Breast MRI for screening:

evaluation of clinical practice and future perspectives

Suzan Vreemann

Uitnodiging

voor het bijwonen van de openbare verdediging van het proefschrift:

Breast MRI for screening: evaluation of clinical practice and future perspectives

Donderdag 27 september 2018 om 10:30 uur in de aula van de Radboud Universiteit Nijmegen Comeniuslaan 2, te Nijmegen

Na afloop van de plechtigheid vindt een receptie plaats in de aula.

Suzan Vreemann suzanvreemann@gmail.com

Paranimfen

Kaman Chung Jan-Jurre Mordang

Breast MRI for screening: evaluation of clinical practice and future perspectives

Suzan Vreemann

Breast MRI for screening: evaluation of clinical practice and future perspectives

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 27 september 2018
om 10:30 uur precies

door

Suzan Vreemann

geboren op 29 juli 1989
te Neede

TABLE OF CONTENTS

Introduction

1

1.1   Breast cancer

Breast cancer is one of the leading causes of cancer death (deaths of any cancer type) in the western world[1,2]. In the Netherlands, one out of 6.6 women faces the diagnosis of breast cancer during her life, assuming a lifespan of 95 years[1,2]. When breast cancer develops, it starts off in the inner lining of the milk ducts or in the lobules that supply these ducts with milk (see Figure 1.1). The two most common types of breast cancer are invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC), which account for 90% of all identified breast cancers[3]. IDC differs from ILC in epidemiology, molecular alterations, clinico-pathological aspects and natural history[4]. Several other (rarer) types of breast cancer exist (medullary, tubular, mucinous, etc.). According to the World Health Organization classification, breast cancer can be classified in up to 21 distinct histological types on the basis of cell morphology, growth and architecture patterns[5]. Moreover, ductal carcinoma in situ (DCIS) is described as a precursor of invasive cancer and is commonly detected in mammography screening, as these tend to be identifiable as calcifications in approximately 25% of cases[6].

In the Netherlands, the general female population between 50 and 75 years of age is currently invited for breast cancer screening with biennial mammography. Screening was introduced in 1988. The regimen is structured in a one-size-fits-all approach, which has shown to be only partly appropriate for breast cancer screening in the entire population because independent patient-related factors like breast density (BD) have a profound effect on the screening accuracy[7].

Today, research focuses on personalizing breast cancer screening programs, for example, by changing towards risk-based screening and adding supplemental screening modalities for women with dense breasts[9,10]. Currently, breast MRI is considered to be the most sensitive imaging method for the early detection of breast cancer. Therefore, women at increased breast cancer risk are recommended to start screening at an early age. In addition, the sensitivity of mammography is relatively poor in these young women due to the large amount of fibroglandular tissue within the breast and consequent masking of breast cancers[11,12]. Since breast MRI is a rather expensive imaging modality, it is only proven to be cost-effective in the highest risk categories. According to the guidelines, these categories include women with a germ-line *BRCA* mutation; other genetic syndromes associated with a high risk of developing breast cancer; or a lifetime risk of $\geq$ 20-25% based upon family history[11,12]. This will be discussed in more detail in Chapter 2.
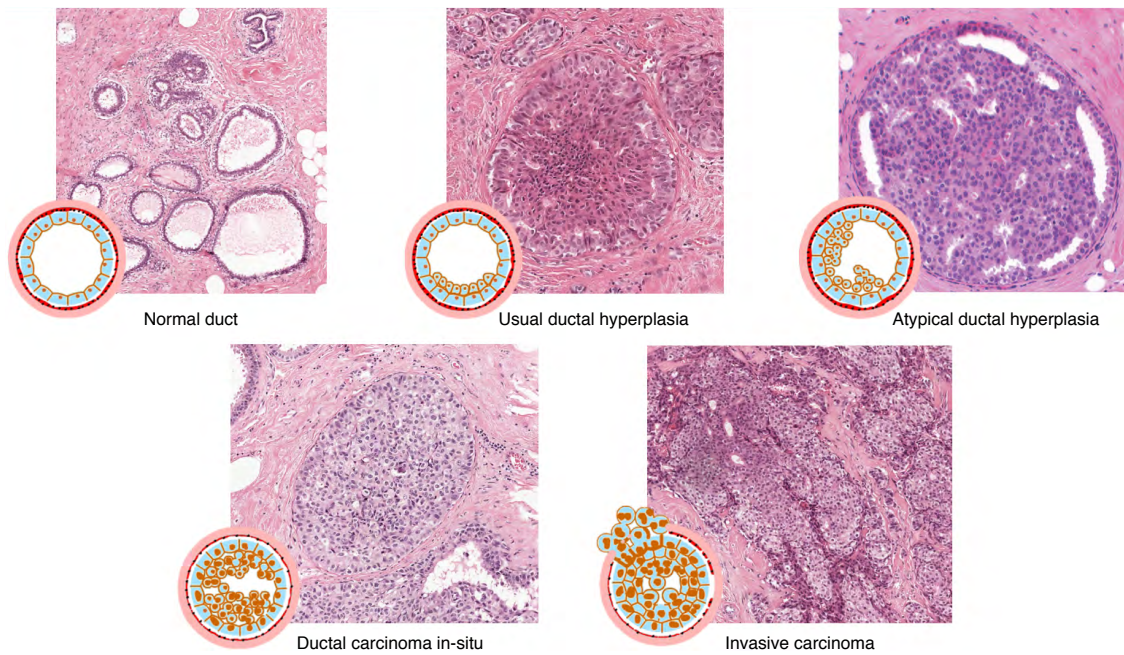
Figure 1.1: Progression of ductal-type breast cancer at different stages. Ductal-type breast cancer is thought to progress from atypical hyperplasia and ductal carcinoma in situ (DCIS) to invasive cancer[8].

1.2   Mammography

Mammography is a fast, relatively inexpensive, widely available, and non-invasive technique yielding a 2D x-ray image of the breast. The x-rays are attenuated differently depending on the composition of the breast tissue they pass. The x-rays are recorded by an x-ray sensitive detector. These attenuation differences among cancerous tissue, glandular, and fatty tissue allow the visualization of breast lesions, which predominantly present as soft tissue masses, grouped calcifications, architectural distortions, or focal asymmetries. The main strengths of mammography are the very high spatial resolution and the possibility to depict microcalcifications that may indicate the presence of in situ or invasive carcinoma[13]. Breast cancer population-based screening using mammography aids in detecting cancers before they become symptomatic. This reduces breast cancer-related mortality by approximately 30%[2]. However, since the attenuation of cancer and fibroglandular tissue is very similar, mammography lacks sensitivity in dense breasts due to tissue overlap and consequent masking of relevant disease (see Figure 1.2).

(a) MLO-view   (b) CC-view
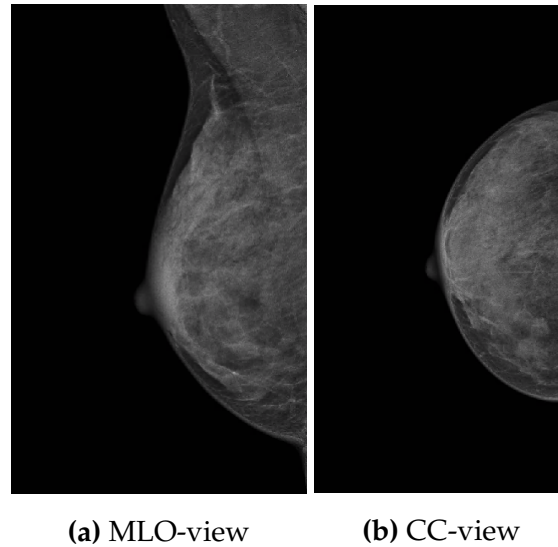
Figure 1.2: Mammography image of the right breast in MLO (a) and CC view (b) of a woman (aged 38) with breast parenchyma of the highest density grade

1.3 Breast MRI

Since the first tests of contrast-enhanced breast MRI with gadolinium diethylenetri-amine penta-acid (Gd-DTPA) were performed in the early eighties, its use has expanded worldwide. Nowadays, dynamic contrast-enhanced breast MRI (DCE-MRI) is widely known as the most effective imaging modality for the early detection of breast cancer[12,14,15]. Initial evaluations showed not only that DCE-MRI was capable of showing most cancers, but also yielded significantly more information than mammography and ultrasound. In prospective studies, a sensitivity of over 90% was documented[15], and more recent reports provide even higher numbers[16,17], likely due to continuously increasing the quality of MRI protocols as well as increased experience with the technique.

The performance of MRI is dependent on the underlying physics. An MRI system consists of several components:

1. A large magnet to generate the magnetic field (B_0). The magnetic field aligns the hydrogen nuclei within the body to generate images.

2. The radio-frequency coil to transmit a radio signal into the body.

3. A receiver coil to detect the radio signals that are emitted by the relaxation of excited nuclei.

4. Gradient coils to provide spatial information of the signals.

5. A computer to reconstruct the received radio-signals into images (by means of Fourier transforms).

The detection of breast cancer in dynamic contrast-enhanced breast MRI is highly improved by the administration of an intravenous contrast agent. This is explained by the fact that tumors are relatively fast-growing and, thus, in need of a lot of nutrients and oxygen. Therefore, they tend to stimulate blood vessel growth. Moreover, the blood vessels that are created by the tumor are leaky, which allow the contrast agent to leak into the surrounding tissue. Thus, the cancerous region will contain more contrast agent, resulting in an increased signal on contrast sensitive sequences. Very indolent lesions may not cause sufficient neovascularization and, thus, contrast leakage to be detected at breast MRI. Moreover, regular breast MRI sequences are not capable of detecting microcalcifications in the breast. This explains why not all lesions detected at mammography are depicted at breast MRI, since MRI, in the case of DCIS, predominantly detects high-grade disease[18].

In the case of enhancing lesions at breast MRI, morphologic evaluation and dynamic series are used to distinguish cancer from benign enhancing lesions. The dynamic series aid evaluation by capturing the peak enhancement and wash-out behavior of the enhancing lesions (see Figure 1.3). The time course of the signal enhancement is of diagnostic value in clinical practice. Morphological (lesion appearance) and kinetic information (enhancement-time curve) of breast lesions are assessed according to the ACR Breast Imaging and Data System (BI-RADS) MRI lexicon[19]. A more detailed elaboration on the current protocols used in breast MRI screening will be presented in Chapter 2.

The high sensitivity of MRI led to the rapid spread of applications and indications for breast MRI. It was recognized as an excellent screening technique as it does not use any ionizing radiation, unlike mammography, and has a very high negative predictive value (the power to predict that there is no breast cancer present). Breast MRI screening is currently only indicated for women at increased risk of breast cancer. International guidelines[11,12] advise women with a lifetime risk of $\geq$ 20 - 25% to be screened with breast MRI and mammography, this includes *BRCA* mutation carriers, women with radiation to the chest at a young age and women with a very high risk due to family history. Women with previously diagnosed high-risk lesions and women with a personal history of breast cancer are also at increased risk and are therefore commonly enrolled in such screening protocols[20]. Due to the many different guidelines available, practice varies from hospital to hospital. Mainly due to the associated costs, the Dutch national guidelines are more restrictive in the population
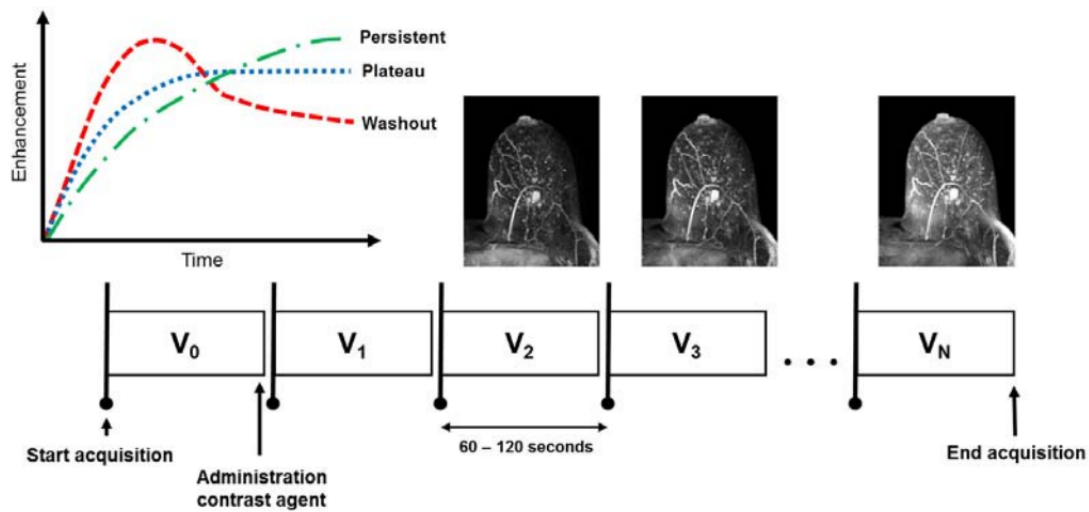
Figure 1.3: Breast MRI acquisition scheme. One volume is acquired before the administration of contrast agent (V_0) and several volumes are acquired after contrast injection (V_1-V_N). With these data, contrast enhancement curves can be computed for a lesion of interest, and lesions can be classified as showing persistent enhancement (green), plateau enhancement (blue), or washout (red)[21].

selected for breast MRI screening and include only women with a lifetime risk of $>$ 50%. In *BRCA* mutation carriers, this screening already starts at the age of 25 with a yearly breast MRI and at the age of 30, a yearly mammography exam is added to this regimen. In women at high familial risk, screening starts at 35 - 45 years, and in women with a personal history of breast cancer, it starts one year after the diagnosis of breast cancer.

1.4   BI-RADS

The Breast Imaging Reporting and Data System (BI-RADS) lexicon is used to report on findings detected in breast imaging. The American College of Radiology developed the first version of this BI-RADS lexicon in 1993 to standardize breast imaging reports, to improve communication with referring physicians, and to provide a quality assurance tool. The lexicon descriptors were designed to predict both benign and malignant disease, eliminate ambiguity, allow automated data collection, and facilitate communication with referring physicians. Structured reports are categorized into several categories, including the description of findings and a final decision-oriented assessment. The first version only included mammography, but currently,

it also includes ultrasound and breast MRI[22]. The final assessment of the exam defined by the BI-RADS lexicon consists of 6 possible categories of which the radiologist needs to choose one, ranging from BI-RADS 1 as being normal to BI-RADS 6 as already proven to include breast cancer. A temporary assessment of BI-RADS 0, meaning that more imaging is required, can also be given.

1.5 Density and background parenchymal enhancement

The BI-RADS lexicon described by the American College of Radiology has a number of standard requirements for the reporting of breast imaging. These include breast density on mammography and background enhancement in breast MRI. Previous research already showed a clear correlation between mammographic breast density and breast cancer risk[7]. However, mammography is a 2D technique and estimating breast density might be inaccurate. Measuring density or the so-called amount of fibroglandular tissue on a 3D technique as breast MRI is more precise, and therefore this might be preferred in for example risk modeling. In addition, case-control studies indicate that background parenchymal enhancement (BPE) on breast MRI might be associated with breast cancer[23,24]. These results show that both breast density measured on mammography and BPE on breast MRI might be potential risk stratification factors for breast cancer.

1.6 Other imaging modalities

As already mentioned, mammography and breast MRI are currently the modalities of choice for high-risk screening, but there are also other potential imaging modalities. First of all, breast tomosynthesis is a relatively new technique that is being implemented into clinical practice and population-based screening programs around the world. For the average risk population, preliminary studies found an increase in sensitivity and an equal or even improved specificity compared to mammography[25]. Tomosynthesis can be described as a special mammography exam that partially overcomes the 2D-limitations (e.g. tissue overlap) by producing a pseudo-3D image of the breast. In tomosynthesis, multiple low dose x-ray projections are obtained at different angles, which are subsequently reconstructed into the final diagnostic images (usually slices of 1 mm thickness). The breast is compressed in the same manner as for mammography, but some studies suggest that a reduction in compression force can be applied in tomosynthesis without losing clinical performance because tissue overlap is less of an issue[26]. Although multiple low dose pro-

jections are acquired, the dose of the exam is comparable or slightly higher than that of a standard mammography exam[27]. Furthermore, there is still debate whether the visibility of calcifications is lower than in mammography[28–30]. Due to the slightly higher dose and the lack of gain in detecting microcalcifications, tomosynthesis is still not used for screening the population at increased risk, and further research is needed to focus on the added value of tomosynthesis when breast MRI is available.

Another modality, which was previously investigated for the implementation of high-risk screening programs, is 3D automated breast ultrasound (ABUS). Ultrasound is a relatively inexpensive, radiation-free, and widely available imaging modality. ABUS is less user-dependent than hand-held ultrasound, which is used in clinical practice, and can be performed by a technician instead of a radiologist, potentially saving time and money. ABUS acquires three-dimensional B-mode US volumes by using a large linear array transducer. Multiple acquisitions are needed to cover the entire breast (usually 2 to 5). The images are then processed to present multiplanar reformats in the coronal and sagittal plane. Especially in women with dense breasts, this technique was speculated to increase cancer detection compared to mammography. This was confirmed by the study of Brem et al.[31]. However, in the study of van Zelst et al.[32] it was shown that, in *BRCA* mutation carriers, ABUS did not lead to the detection of additional cancers beyond that available with mammography and breast MRI, while it did lead to more false positive findings. Similar findings were reported for supplemental hand-held ultrasound[15]. Consequently, the performance of currently available ultrasound techniques for screening the high-risk population that is also screened with breast MRI is not recommended. Van Zelst and coworkers discussed the potential cause of this limited value of ultrasound in high-risk screening explaining that this might be caused by the relatively benign appearance of the typical *BRCA*-related cancers.

1.7 Quality control of screening programs

Quality control in screening programs mainly uses defined performance measures, described by Roseberg et al.[33]. These performance measures include recall rate, biopsy rate, positive predictive value of recall, positive predictive value of biopsy, cancer detection rate, sensitivity, and specificity. These measures have been used in Chapter 3 and can be explained as:

- **Recall rate (RR)**, which is described as the number of recalls per 1000 screening examinations.

- **Biopsy rate (BR)**, which is described as the number of biopsies performed per 1000 screening examinations.

- **Positive predictive value of recall (PPV1)** is calculated by dividing the number of correct recalls (resulting in cancer diagnosis) by the total number of recalls.

- **Positive predictive value of biopsy (PPV3)** is calculated by dividing the number of correct biopsies (resulting in cancer diagnosis) by the total number of biopsies.

- **Cancer detection rate (CDR)**, which is described as the number of breast cancers detected per 1000 screening examinations.

- **Sensitivity**, which is described as the number of true positives (breast cancers detected in screening examinations) divided by the total number of positives. Positives can be divided into true positives and false negatives, where false negatives are the examinations that were reported to be negative but a cancer was subsequently found.

- **Specificity**, which is described as the number of true negatives (screening examinations that not include a breast cancer, checked with the proper follow up) divided by the number of negatives. Negatives can be divided into true negatives and false positives, where false positives are described as the examinations that were reported to be suspicious for breast cancers, but no breast cancer was detected after appropriate follow-up.

Guidelines describe accepted and desired levels for all these performance measures for mammography screening[34]. Recently, auditing guidelines were also described for breast MRI[19], however, evaluation of prior examinations was not mentioned. Although these quality controls/audits are regularly performed in mammography screening, in breast MRI these are rare. When reported, these are mainly reported in prospective trials with sharply defined population groups. At present, quality control results for breast MRI screening programs on a hospital level do not exist.

1.8 Thesis outline

This thesis aims to evaluate and optimize breast MRI screening practice, and analyses the performance of a breast MRI screening program at a single academic hospital. For several chapters (**Chapter 3 to 5**) the audit practice of the Dutch national screening program for breast cancer was used as a blueprint. Subsequent chapters analyze

methods to optimize screening practice by including background parenchymal enhancement into the risk stratification. Finally, the last study investigated a novel MRI sequence, developed for optimizing screening practice.

In detail, **Chapter 2** provides an introduction to breast MRI screening and the protocols used for breast MRI screening. In **Chapter 3** the performance of the current high-risk screening program at our institution was assessed, providing detailed performance measures for women in different risk categories. **Chapter 4** subsequently describes, within the same population, the (lack of) added value of mammography when breast MRI is available. In **Chapter 5**, we performed a more in-depth audit of cancers detected in the high-risk screening program, investigating the visibility of these cancers on prior MRI scans that were reported to be normal. In **Chapter 6**, we started exploring the possibility to use BPE as a risk stratifier for breast cancer, evaluating whether the assessed BPE is independent of tumor type. In **Chapter 7**, we subsequently investigated, in a prognostic cohort study, whether BPE is indeed an independent risk factor for breast cancer. Finally, **Chapter 8** introduces different approaches for increasing the cost-effectiveness of breast MRI and introduces **Chapter 9**, which discusses the potential of a new MRI sequence for breast MRI screening, possibly increasing patient comfort and cost-effectiveness, while still providing excellent diagnostic capacity. A general discussion regarding the studies in this thesis is provided in **Chapter 10**, while **Chapter 11** and **Chapter 12** provide summaries in both English and Dutch.

MRI protocols for breast cancer screening

2

Abstract

Magnetic resonance imaging (MRI) techniques for breast cancer screening keep evolving. At present, guidelines state only minimal requirements for breast MRI screening protocols. There is still no consensus on the optimal protocol to be used for breast MRI screening. New breast MRI screening sequences and shorter protocols are introduced to optimize breast cancer detection and to minimize false-positive recalls. In this chapter, recent advances in breast MRI protocols and sequences are discussed to get a better understanding of what is possible and what is needed in breast MRI for screening. We discuss the current state-of-the-art of breast MRI for screening and evaluate the requirements for basic sequences. Subsequently, novel approaches for breast MRI screening that are currently on the brink of clinical implementation are presented. Finally, techniques that are now only used for research purposes, but yield promise for future implementation are discussed.

2.1   Introduction

Magnetic resonance imaging (MRI) is nowadays widely known as the most effective imaging modality for the early detection of breast cancer. Contrast-enhanced breast MRI has been increasingly used since the Eighties when an excellent breast lesion conspicuity after intravenous injection of gadopentetate dimeglumine (Gd-DTPA) was firstly shown[35]. Thanks to the acquisition of multiple series of contrast-enhanced images, the method rapidly evolved into dynamic contrast-enhanced (DCE) breast MRI[36]. Initial evaluations showed not only that DCE-MRI was capable of showing most cancers, but also yielded significant additional information over mammography and ultrasound. In subsequent studies, a sensitivity of over 90% was documented for breast malignancies.

The high sensitivity led to the rapid spread of applications and indications for breast MRI. It was thought to be an excellent screening technique, as it does not use any ionizing radiation and has a very high negative predictive value. A standard protocol for DCE-MRI was adopted, consisting of at least one T1-weighted acquisition before contrast injection and several acquisitions after contrast injection. Other contrast materials than the original Gd-DTPA were occasionally used.

Multiple prospective trials investigated the value of DCE-MRI using variations of this protocol for breast cancer screening in women at various classes of increased risk and reported a sensitivity in the range of 71% to 91% (see Table 2.1).

Using the results of early studies, skeptics pointed at the so-called low specificity of breast MRI as a flaw that prevented the use of MRI in screening. Since in screening most scans are normal, a high number of false positives would lead to a very high recall rate, and this would pose an unacceptable burden on healthy women and large logistic issues to health care systems. This led to a wide range of additional acquisitions, including not only old and new T2-weighted sequences but also diffusion-weighted imaging (DWI), and even proton spectroscopy, each of which was shown to have a potential for reducing false positives of breast MRI. Therefore, the current state-of-the-art protocols in breast MRI are multi-parametric in nature. Nonetheless, the so-called low specificity of breast MRI is a complex questionable phenomenon to which the readers inexperience and the absence of prior screening MR studies also gave critical contributions.

Even though the performance of MRI screening with multi-parametric protocols

Table 2.1: Sensitivities and specificities for breast MRI screening of women with an elevated breast cancer risk

Author[ref.]	Year	No. of women screened	No. of cancers	Sensitivity (%)	Specificity (%)
Kriege[14]	2004	1909	45	71	90
Leach[37]	2005	649	35	77	81
Kuhl[38]	2005	529	43	91	97
Rijnsburger[39]	2010	2157	75	71	90
Trop[40]	2010	184	12	83	94
Sardanelli[41]	2011	501	52	91	97
Passaperuma[42]	2012	496	57	86	90
Riedl[15]	2015	559	40	90	89

seems excellent, the scan-time required, as well as the time needed for evaluation, and thus the associated costs, are high, and this currently contributes to limit the use of breast MRI as a screening tool in high-risk women.

Current research, therefore, focuses on the development of shorter imaging protocols to reduce scan-times and to cut costs. In addition, imaging biomarkers are extracted focusing on the evaluation of aggressiveness of breast cancer, in order to open a way for characterizing those cancers that are biologically active. Finally, new techniques that aim at excluding contrast agent administration from the acquisition are under development. Updates on the recent developments on these topics will be discussed in this chapter.

2.2   Indications for breast MRI screening

The risk-level of women included in various MRI screening studies is highly heterogeneous. The MRISC trial[14] included all women with a lifetime risk of 15% or more, whereas for example the Canadian trials[42] only included women with *BRCA* mutations and their first-degree relatives. In 2007, the American Cancer Society (ACS) published a guideline[12] on screening with supplemental MRI. They concluded that MRI as an adjunct to mammography was indicated for all women with a lifetime risk of 20 - 25% as assessed with BRCAPRO or similar tools that mainly focus on family history. In 2008, these guidelines were also adopted by the European society for breast imaging (EUSOBI), as the recommendations were mainly based on European

studies[11]. The group of high-risk women considered by these guidelines consists of women with either a proven *BRCA1* or *BRCA2* germline mutation, other genetic syndromes associated with a high incidence of breast cancer, or a $\geq$ 20 - 25% lifetime risk of developing breast cancer. Women with a history of chest radiation therapy between the ages of 10 and 30 also have a similar risk and should be screened with supplemental MRI.

For women at intermediate risk (defined as 15 - 20% lifetime risk, or a personal history of invasive or in-situ carcinoma, lobular neoplasia, or atypical hyperplasia) the use of breast MRI for screening is still under investigation, and there has been no clear statement yet. It is agreed on that breast MRI screening should currently not be used in women with a lifetime risk lower than 15%[43]. However, screening trials with MRI for women at average risk, but with very dense breast are on their way, thus introducing risk factors for stratification that were not used in previous screening studies or risk models. Whether or not screening guidelines should be expanded to also include these women will become evident in the upcoming years.

Despite the above-mentioned guidelines, national guidelines are often more conservative as they focus more on cost-effectiveness. While there is a clear increase in breast cancer detection in all risk groups, cost-effectiveness is only proven for women at the highest risk (e.g. those with a *BRCA* mutation and their first degree relatives, and those with a history of chest radiation therapy before 30), and hence several current national breast cancer screening guidelines only state that a yearly breast MRI is appropriate for women at very high risk for developing breast cancer.

2.3   Requirements for breast MRI

The European Society of Breast Imaging (EUSOBI)[11], the European Society of Breast Cancer Specialists (EUSOMA)[44], and the American College of Radiology (ACR)[45] have specified requirements for the performance of breast MRI. So far, these requirements are general and not specifically focused on breast MRI for screening. In light of the current diversion of MRI for screening, MRI for staging, and other indications, these guidelines will likely be adapted in the near future to be more specific for the indication for which breast MRI was performed. However, the minimal requirements as specified by these organizations remain vital as they are very liberal. All breast MRI protocols should include T1-weighted pre- and post-contrast sequences to report on lesion morphology and enhancement features with sufficient spatial and temporal resolution. All other sequences are deemed supplemental, however,

the use of additional T2-weighted acquisitions is generally endorsed.

Minimal requirements are as follows[11,44]:

- Field strength
  A 1.5T magnet is considered a minimum technical requirement because of the relationship between field strength and resolution (the advantages and disadvantages of 3.0T imaging will be discussed later in this chapter).

- Spatial resolution
  High spatial and temporal resolution are needed to detect and characterize small abnormalities. The EUSOBI guidelines state that the slice thickness should not be higher than 2.5 mm and the in-plane resolution should be 1 mm^2 or less, thus, minimizing the problem of volume averaging effects. Other guidelines still accept up to a 3 mm slice thickness.

- Scan plane
  No absolute preference for scan plane is recommended by guidelines. While in past times some technical issues favored the use of coronal planes with the aim of optimizing both temporal and spatial resolution, currently axial and sagittal planes are preferred, also for the evaluation of symmetry.

- Fat saturation and temporal subtraction
  Spectral fat saturation can be used to reduce the fat signal while preserving the signal-to-noise ratio (SNR), however, it is not mandatory. The guidelines state that radiologists must not solely rely on temporal subtraction images for the assessment of enhancement, since this may result in misregistration due to patient motion. When motion artefacts do appear, motion correction might be helpful in reducing artefacts encountered with image registration.

- Radiofrequency coils and simultaneous bilateral imaging
  Guidelines state that simultaneous bilateral high-resolution images should be acquired as breasts are symmetric organs and comparison between the two breasts can be performed. The use of a multichannel dedicated bilateral breast coil is mandatory[38,40,46,47]. These are commercially available and provide excellent spatial and temporal resolution for improved visualization of small lesions[48].

- Contrast agent
  Breast MRI without contrast is not acceptable according to current standards,

except for the evaluation of implant integrity. Therefore, all screening examinations should be contrast-enhanced studies. As a contrast agent, a two-compartment (vascular / interstitial) gadolinium-chelate should be administered intravenously as a bolus with the standard dose of 0.1 mmol/kg with an injection rate of 2 - 3 ml/s, followed by saline flushing (20 - 30 ml at 2 ml/s), preferably using an automatic injector.

- Temporal resolution
  The time interval between images series for DCE studies should be no longer than 120 s[44].

- Volume of MRI studies per institution
  The EUSOMA[44] recommends that a minimum number of 150 cases needs to be performed per institution per year. Despite the lack of recommendations from the other societies, this seems a wise recommendation, especially when considering the screening setting. Even in very high-risk screening, the cancer detection rate is only in the order of 2 - 3%, and hence at a rate of 150 screening examinations, only 3 cancers are detected per year. Therefore, some centralization and significantly larger volumes are recommended.

2.4 MRI sequences

2.4.1 T1-weighted sequences

To understand the basics of sequences used for screening, some knowledge of MRI physics is desirable. One of the most common pulse sequences in MRI is the T1-weighted sequence, also referred to as the spin-lattice relaxation sequence. The images obtained from this sequence display the differences in T1 relaxation times of different tissues. This sequence relies upon the longitudinal relaxation of the tissues net magnetization vector (T1 relaxation describes the spin relaxation in the z-direction). T1-weighting is achieved with short echo times and repetition times. As fat quickly realigns its longitudinal magnetization with $B0$ (the main magnetic field), it appears bright on a T1-weighted image. Conversely, water has a much slower longitudinal magnetization realignment after a radiofrequency pulse, and therefore, has less transverse magnetization. Thus, water has a low signal and appears dark. Figure 2.1 shows examples of the different images obtained using a T1-weighted sequence either with or without fat suppression. Gadolinium-based contrast media have a paramagnetic effect on the tissue and hence reduce the T1 relaxation time (i.e., the time needed for longitudinal relaxation). This increases the signal of the tissue and

hence, a high signal (that appears bright on the image) is produced in areas of contrast agent uptake[49].

In breast MRI screening protocols, T1-weighted sequences are, thus, used for detection of areas where the contrast agent accumulates, such as malignant breast lesions[11,45]. Several T1-weighted sequences are obtained in dynamic succession to visualize the course of tissue contrast-enhancement.

In the early years of breast MRI, it was necessary to choose between either temporal or spatial resolution. Two fundamentally different protocol designs evolved, the static design and the dynamic design. The static design was most popular in the United States and specifically evaluated the morphologic features of enhancing lesions at high spatial resolution. The dynamic design was instead mostly favored in European countries with the aim of using dynamic enhancement characteristics to distinguish benign lesions from malignant lesions[50]. Nowadays, thanks to the technical progress that has been made, it is possible, to a certain extent, to integrate these demands, and the final evaluation is virtually always based upon a combination of morphologic and dynamic enhancement features[50].

The BI-RADS MRI-lexicon[51] states the different enhancement patterns that have to be rated based on T1-weighted sequences. All findings should be viewed on both pre- and post-contrast scans and both morphologic and kinetic characteristics should be evaluated. Malignant lesions tend to enhance rapidly, typically reaching 90% of peak enhancement within 60 seconds following injection, while fibroadenomas and other benign lesions tend to enhance at a lower rate. Strong early enhancement with a relative signal increase of over 140% and a peak of enhancement before 3 minutes together with an early washout (signal decrease of more than 10% following maximum enhancement)[46] is highly suggestive of malignancy.

Different T1-weighted sequences exist, as shown in Table 2.2. Spin-echo sequences are generally not recommended as these are too slow to achieve the spatial and temporal resolution required for breast MRI screening. The EUSOBI recommends at least a T1-weighted spoiled gradient echo (GE) pulse sequence before the administration of contrast agent, one at peak enhancement approximately 90 seconds after contrast agent administration, and one 5 to 7 minutes after contrast agent administration in order to investigate the morphology and the dynamics of enhancement. Most protocols that are currently in use include one pre-contrast T1-sequence and 3 to 5 post-contrast T1-sequences for the dynamic evaluation[11,50,52–55].
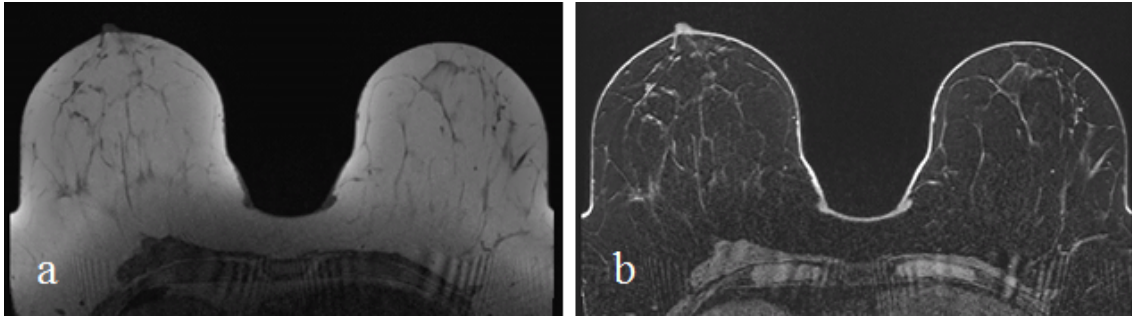
Figure 2.1: Axial T1-weighted images in a 42-years-old *BRCA2* mutation carrier with relatively fatty breast, using a sequence without fat suppression (a) or with fat suppression (b).

In order to obtain a more reliable T1-weighting, GE sequences used for DCE breast imaging use the so-called spoiler (typically a radiofrequency spoiler) that disrupts transverse coherences that may persist from cycle to cycle of the sequence. Thus, immediately before each radiofrequency pulse, the steady-state magnetization has no transverse components, while the longitudinal magnetization reaches a steady-state[56].

However, spoiled gradient-echo sequences exist in both two-dimensional (2D) and three-dimensional (3D) acquisition modes. It is still unknown which of the two is the best, or most appropriate one for breast MRI, in particular for screening. Both methods have their own advantages and disadvantages. When comparing 2D to 3D sequences, 3D sequences are known for their higher T1-contrast and higher SNR, resulting from shorter repetition times and echo-times. The higher SNR can be used to improve the spatial resolution, both in-plane (pixel size) and through-plane (partition thickness). However, 3D imaging may suffer from image degradation (pulsation-, susceptibility-, and ghosting artefacts) and therefore, some prefer 2D imaging[56].

2.4.2 T2-weighted sequences

Most of the current breast MRI screening protocols include a T2-weighted sequence (examples are presented in Figure 2.2). In the EUSOBI guidelines for breast MRI, it is stated that this sequence can be useful in the differentiation between benign and malignant lesions (and, thus, increasing specificity and positive predictive value), as in most cases cancer does not yield a high signal on turbo spin-echo T2-weighted images without fat suppression, whereas many benign lesions do. However, as most

Table 2.2: T1-weighted sequences

Sequence	Characteristics
Spin-echo (SE)	• T1-weighting is maximized by setting repetition time (TR) to be similar to or slightly shorter than the T1 values of the tissue of interest, while setting echo-time (TE) as short as possible • It takes time to rephase the signal, thus increasing the scan time
Turbo spin-echo (TSE)	• Because of the long scan time of SE sequences, TSE meets the demand for faster imaging. Multiple echoes are formed and measured by adding several 180° pulses and measuring a spin-echo after each pulse, rather than measuring a single echo after each 90°-180° combination
Fast advanced spin-echo (FASE)	• Half Fourier imaging is used, which shortens scan time even more than TSE
Gradient-echo (GE, GRE)	• Similar to spin-echo sequences with two main differences: 1. Use of an initial 70° pulse, rather than a 90° pulse 2. Use of gradient reversal, instead of a 180° pulse to form an echo, thus eliminating the need to wait to allow a regrowth of the longitudinal magnetization • T1-weighting is achieved by using a short TR, very short TE, and a moderate flip angle • Due to higher T1 contrast and shorter acquisition times, these sequences are generally preferred over spin-echo sequences and are therefore recommended by guidelines

of these lesions can also be identified in T1-weighted images, the EUSOBI guidelines state that there is no clear evidence of the added value of T2-weighted sequences in screening yet[11].

In 1999, Christiane Kuhl and coworkers were the first investigators to evaluate the added value of T2-weighted imaging for breast MRI[57]. They investigated whether T2-weighted pulse sequences can help in the differential diagnosis of enhancing lesions on dynamic breast MRI. Fibroadenomas and well-circumscribed breast cancers may have a similar appearance, as both may present as a rapidly and strongly enhancing focal lesion. Fibroadenomas and breast cancers tend to demonstrate different signal intensities on T2-weighted imaging. In practice, if a well-circumscribed enhancing lesion is detected in breast MRI, high signal intensity in the corresponding T2-weighted image can be used to support the diagnosis of a benign lesion. In the higher age groups (over 40), a low T2-weighted signal should arouse suspicion of malignancy, even though the lesion is well circumscribed. The authors concluded that T2-weighted turbo spin-echo sequences can be helpful as an adjunct to the dynamic breast MRI protocol, and it should not be used as a stand-alone approach but in conjunction with and secondary to criteria like enhancement kinetics and morphological data. Laura Heacock and coworkers[54] came to similar conclusions. In their study, the addition of a T2-weighted sequence to the dynamic protocol resulted in a higher lesion conspicuity but had no effect on cancer detection. Unfortunately, the effect of the T2-weighted images on specificity was not evaluated.

The use of spectral fat-saturation as added to T2-weighted sequences was specifically evaluated in the above-mentioned work by Kuhl and coworkers[57]. They stated that the turbo spin-echo (TSE) pulse sequence without fat suppression is the only suitable sequence to assist in lesion characterization. If used cautiously, T2-weighted imaging should improve the diagnostic accuracy of diagnostic breast MRI by helping to avoid false-positive diagnoses, particularly in young women. Therefore, a T2-weighted turbo (also called fast) spin-echo sequence can be performed as a start of a breast MRI screening, before the dynamic protocol. In combination with fat-saturation, this sequence can also be used to identify cysts, as cysts have extremely long T1 and T2 values relative to other breast tissues. Cysts typically have a few macromolecules to shorten T1 and lack of cellular structure to shorten T2. Thus, cysts appear darker on T1-weighted sequences, while they appear much brighter than other tissues on T2-weighted sequences due to their longer T2-values and higher hydrogen densities. Hence, cysts are easily identifiable on fat-suppressed T2-weighted imaging[58]. However, the necessity to detect cysts in a screening protocol remains questionable and
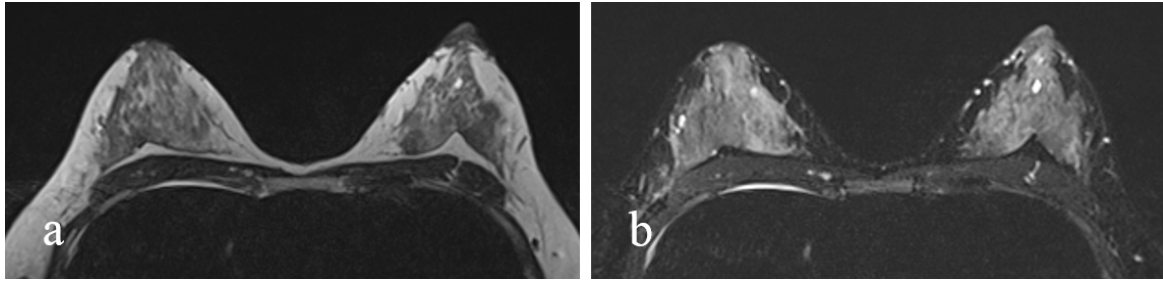
Figure 2.2: Axial T2-weighted image in a 38-years-old woman at familial risk of breast cancer with dense breasts, using a sequence without fat suppression (a) or with fat suppression (b).

this alone should not be used as an argument to perform a T2-weighted acquisition.

2.4.3 Diffusion-weighted imaging (DWI) and other additional sequences

DWI is sensitive to the mobility of water molecules in tissue. Water motion is most commonly quantified by means of an apparent diffusion coefficient (ADC), a model using the principle that tissue-confined water behaves similarly to free water, but with reduced diffusivity (see Figure 2.3). This is used as a marker of cellularity in oncologic imaging, as it represents a decrease in extracellular space relative to the more viscous intracellular fluid of proliferating cells. Measurement of diffusivity does not require application of contrast agents. Thus, DWI might be a suitable technique for non-contrast breast MRI, as we will discuss below.

Notably, the signal obtained by DWI sequences is based on a T2-weighted acquisition. Diffusivity is measured by applying a strong spoiler gradient to excited hydrogen protons. After a certain period of time, this spoiler gradient is reversed. Strength and duration of the spoiler gradient together determine the so-called b-value: the higher the b-value, the stronger the signal positive correlation with tissue diffusivity. In fact, protons that do not move regain their signal, whereas moving protons experience a different gradient at both instances and hence lose their signal[59]. Some studies[60,61] focused on the choice of b-values for DWI of the breast. A b-value near to zero (or 50 to reduce signal from vessels) and a b-value around 800-850 are sufficient for clinical imaging. However, for more elaborate DWI techniques such as diffusion tensor imaging (DTI) or intra-voxel incoherent motion (IVIM) imaging, multiple and higher b-values are required.

Studies have shown the potential of DWI to increase breast MRI specificity. As

discussed by Gurpreet S. Dhillon and coworkers[48], DWI has a higher specificity to differentiate benign from malignant lesions than contrast-enhanced MRI. However, to not lose in terms of sensitivity, a feasible way to implement DWI in a multi-parametric protocol appears to adapt BI-RADS scores based on ADC values as proposed by Katja Pinker and coworkers[62]. Differences in ADC values may be able to distinguish ductal carcinoma in situ (DCIS) from both normal tissue and invasive ductal carcinoma. The ADC value is lowered in DCIS compared to normal breast parenchyma, but is still significantly higher than the ADC values seen in invasive ductal carcinoma. However, this intermediate ADC value is not specific to DCIS and might overlap with other benign and malignant lesions[63].

Several studies investigate the added value of DWI when DCE-MRI is available. Sibel Kul and coworkers[64] applied that this strategy to 84 breast lesions, showing that the combination of DWI and DCE-MRI had the potential to increase the specificity of breast MRI, a result which was confirmed by Richa Bansal and coworkers[65] in a larger study including 232 lesions.

Other sequences have been proposed to further improve the specificity, including dynamic susceptibility contrast (DSC) imaging, which is a kind of perfusion imaging based on the T2* effect of contrast agent (contrast uptake reduces the T2* value, thus resulting in a lowered signal) and susceptibility-weighted imaging (SWI), which is based on the tissue inhomogeneities and calcifications (again resulting in a lowered signal on T2*-weighted images)[66,67], as well as various types of mainly 1H-based magnetic resonance spectroscopy (MRS) approaches[68–70]. Although each of the techniques has some merits, none of these approaches really made it into clinical practice. Especially for screening purposes, they are currently obsolete. Faster approaches and techniques to increase the signal might, however, in the future render these techniques viable again.

A recent comparison between the evidence in favor of DWI and that in favor of spectroscopy found that DWI is certainly the winner[71]. As mentioned above, DWI, considering both the robustness and short acquisition times, entered breast MRI clinical protocols as the most used additional sequence.

2.5 Breast MRI at 3.0T: advantages and disadvantages

As 3.0T systems become more widely available throughout the world, many facilities may consider performing breast MRI, including screening protocols, at this higher
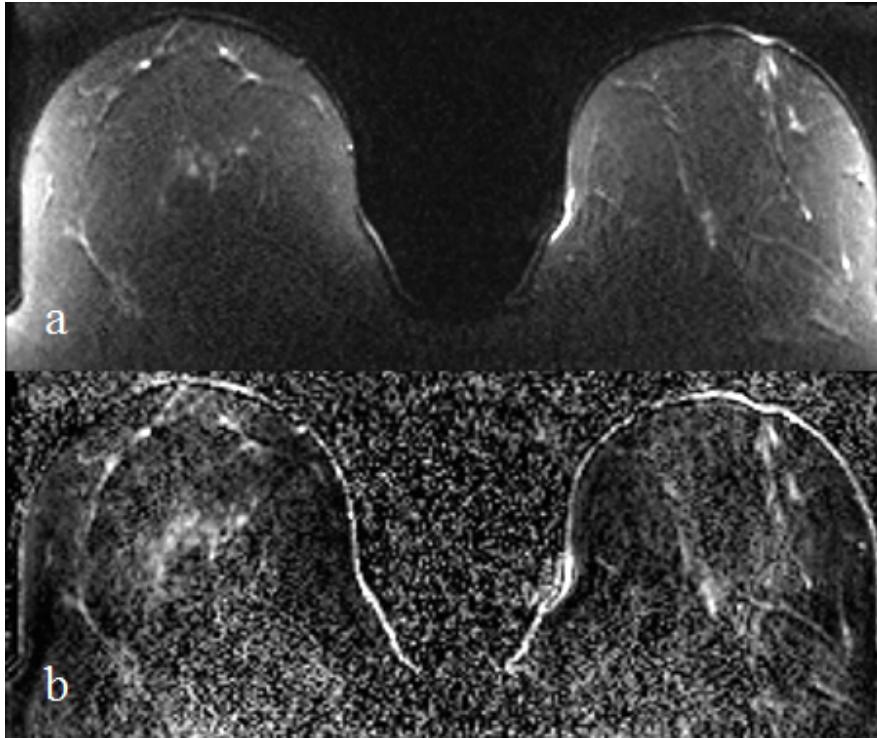
Figure 2.3: Axial diffusion-weighted images in a 42-years-old *BRCA2* mutation carrier using DWI (b=850 (a) and the corresponding ADC map (b).

field. In fact, it is widely known that a higher field strength results in a higher SNR. The improved SNR should, in theory, allow for a better visualization and characterization of enhancing lesions, which may improve the detection of breast cancers[45]. However, this higher field strength also causes an increased field inhomogeneity, which is a clear disadvantage. In the case of 3.0T imaging, several image artefacts can be categorized according to their main underlying mechanism, such as increased SNR, susceptibility variation, chemical shift, or decreased radiofrequency wavelength[72,73].

Increased SNR can result in more pronounced Gibbs ringing artefacts at 3.0T compared to 1.5T. These artefacts occur when Fourier transforms are used to reconstruct MR signals into images. Any signal (and thus every image) can be represented as an infinite summation of sine waves of different amplitudes, phases, and frequencies. In MRI, we sample a finite number of frequencies and we approximate the image by using relatively few sine waves in its Fourier representation. In other words, the Fourier series is cut short. Gibbs ringing artefacts are prominent at high-contrast interfaces, manifested by variable undershoot and overshoot oscillations. These artefacts can have a variety of forms, including false widening of edges, enhancement of the edges, or distortion of tissues.

Ghosting artefacts are associated with parallel imaging (commonly used in MRI with the aim to decrease acquisition times at both 1.5T and 3.0T) and are usually more severe at 3.0T. In fact, the increased SNR at 3.0T, especially in combination with a high channel count (8 or more coils, which is currently standard of care in most clinical practices) used for parallel imaging can worsen the problem.

Chemical shift artefacts are more often present at 3.0T. At a fixed receiver bandwidth, the fat-water chemical shift will be twice as many pixels compared to 1.5T.

Susceptibility variations due to the presence of implanted foreign bodies can also cause local non-uniformity of the main magnetic field resulting in several artefacts, including non-planar 2D slices, in-plane image distortion, and local regions of hypo- and hyperintensity. These artefacts are also stronger at 3.0T. Sequences with long echo trains suffer the most from susceptibility variations. To decrease these artefacts, parallel imaging could be used, taking the comments on the above-mentioned ghosting artefacts into account.

Because 3.0T imaging is already well introduced in clinical imaging, most of these artefacts can be overcome by several already available methods[72]. Clinical 3.0T images are adequate in general, even though it remains uncertain whether these are better than 1.5T images.

There are only a few comparisons between breast imaging at 1.5 and 3.0T. Christiane Kuhl and coworkers[74] described intra-individual results in 37 women with 53 lesions. At 3.0T, the image quality was slightly but significantly better and the diagnostic confidence as measured at receiver operating characteristics analysis was significantly higher. No susceptibility effects were observed. Motion artefacts were observed at 3.0T and 1.5T at the same rate and degree. The smaller pixel size in patients who were examined at 3.0T with high in-plane imaging matrices could result in subtraction artefacts due to motion and therefore, degraded the image quality. Nevertheless, the higher spatial resolution at 3.0T helped to improve classification of 11 of 51 lesions (2 were excluded because of insufficient enhancement).

Ana P. Lourenco and coworkers[47] reported on a comparison of 495 3.0T versus 650 1.5T breast MRI screening scans. They found a significant increase in both biopsy recommendation rate and the positive predictive value of biopsy at 3.0T. Notably, cancer detection rate was significantly higher at 3.0T (2.6%), compared with 1.5T (0.9%).

These results, even limited by the retrospective inter-individual design, showed a potential for a greater efficacy of breast MRI screening at 3.0 T.

Thus, despite existing disadvantages from scanning at a higher field strength, 3.0T breast MRI still improves the diagnostic confidence and the cancer detection rate in a screening population. However, 1.5T breast MRI remains adequate in most settings when updated protocols and breast coils are used.

2.6 Screening sequence protocols

Examples of breast MRI screening sequence protocols screening protocols used at 1.0T, 1.5T, and 3.0T are reported in Table 2.3. However, we should consider that a large variety of technical options were used, although all of them were mainly based on a 2D or 3D spoiled gradient-echo dynamic series.

2.7 Abbreviated screening protocols

While breast MRI screening, as currently implemented, has shown great potential for early detection of cancer in women at increased risk of breast cancer, its wide implementation remains difficult. This is largely caused by the high costs of MRI itself. In addition, the huge amount of image series produced lengthens the reading time and makes actual mass screening very difficult.

Therefore, several research groups focused on reducing the time required for scanning and evaluating breast MRI. Evidence is mounting that shorter protocols, in fact, are just as good for screening as the much lengthier multiparametric protocols that are currently in use. This could potentially increase the access to breast MRI by significantly reducing the cost and time associated with the examination, both the acquisition time and the radiologists reading time.

Several abbreviated protocols were described by different groups of authors[52–54,76] (see Table 2.4). In 2014, Christiane Kuhl and coworkers[76] presented the first and simplest version of abbreviated breast MRI. The protocol is condensed into one pre-contrast and one post-contrast acquisition. The subsequent generation of subtraction images and maximum intensity projections (MIPs) also renders reading exceptionally fast. Reading time of the MIPs was reported to be below 2 seconds. In their study, it was evaluated whether this abbreviated protocol would be sufficient to

Table 2.3: Examples of screening protocols

First author [ref.#]	Breast MRI protocol		Sequence parameters	
Sardanelli [41]*	1.0 T or 1.5T			
	1.	3D T1-weighted spoiled gradient echo (1 pre-contrast and 5 post-contrast)	1.	Transverse or coronal plane; TR: 13ms, flip angle: 20-30°; partition thickness: $\leq$3mm, number of partitions: 40-128; acquisition time up to 120s; pixel size up to 1.4 x 1.4 mm
Kuhl [75]	1.5T			
	1.	2D or 3D T-weighted spoiled gradient echo series (1 pre- and $\geq$ 4 post-contrast)	1.	Spatial resolution $\leq$1x1mm (in-plane), $\leq$ 3 mm (through-plane); acquisition time 120s
	2.	T2-weighted sequence	2.	Spatial resolution $\leq$1x1mm (in-plane), $\leq$ 3 mm (through-plane) Total acquisition time 15 min
Emaus [9]	3T			
	1.	T2-weighted sequence (optional)	1.	Acquisition VS $\leq$ 1.11 x $\leq$ 1.46 x $\leq$ 4.00 mm Reconstruction VS $\leq$ 0.90 x $\leq$ 0.90 x $\leq$ 4.00 mm Acquisition time 147-248s
	2.	DWI sequence	2.	Acquisition VS $\leq$ 2.25x $\leq$ 2.51 x $\leq$ 5.00 mm Reconstruction VS $\leq$ 1.55 x $\leq$ 1.70 x $\leq$ 4.00 mm Acquisition time 215-301s, b-values 0, 50 or 150, and 800 s/mm^2
	3.	DCE data sets		
	(a)	High spatial resolution pre-contrast	(a)	Acquisition VS $\leq$ 1.00 x $\leq$ 1.00 x $\leq$ 2.00 mm Reconstruction VS $\leq$ 1.00 x $\leq$ 0.94 x $\leq$ 1.00 mm Acquisition time 80-152s
	(b)	High temporal resolution series before and during the first seconds after contrast injection	(b)	Acquisition VS $\leq$ 2.58 x $\leq$ 2.82 x $\leq$ 6.00 mm Reconstruction VS $\leq$ 1.18 x $\leq$ 1.18 x $\leq$ 3.00 mm Acquisition time 147-248s, pre-contrast acquisitions (N=1), post-contrast acquisitions (N=15-19)
	(c)	High spatial resolution series	(c)	Acquisition VS $\leq$ 0.90 x $\leq$ 1.00 x $\leq$ 1.80 mm Reconstruction VS $\leq$ 1.00 x $\leq$ 0.80 x $\leq$ 1.00 mm Acquisition time 147-248s, acquisitions (N=5-6)

VS; voxel size

* Although not considered in the final report [41] (the BI-RADS classification was mainly based on DCE-imaging), a T2-weighted sequence was also included in the protocol.

identify breast cancer in a screening cohort. The full diagnostic protocol included a T1-weighted pre-contrast and 5 post-contrast scans followed by a T2-weighted sequence and a coronal T1-weighted sequence. While the full protocol needed about 17 min, the abbreviated protocol needed only 184 seconds. The overall sensitivity of the abbreviated protocol was 100% (negative predictive value 99.8%) with a specificity of 94.3%. However, only 11 cancers were detected overall. With the use of the full diagnostic protocol the characterization of findings classified as possibly benign (BI-RADS 3) was improved, showing that the additional pulse sequences in the full protocol are mainly needed for lesion characterization.

In 2015, Victoria L. Mango and coworkers[52] looked into the sensitivities per sequence of the abbreviated protocol. They found a mean sensitivity of cancer detection of the first post-contrast sequence of 96%, equal to the first post-contrast subtracted sequence. Sensitivity using only the MIPs was significantly inferior (93%), which must be taken into account when deciding to screen using only MIPs.

In the same year, Lars J. Grimm and coworkers[53] tested two different abbreviated protocols in a specifically designed case series of 48 patients selected from high-risk screening. One protocol consisted of a T2-weighted sequence, as well as the pre-contrast and the first post-contrast T1-weighted sequences. In the other protocol, the second post-contrast T1-weighted sequence was added to the sequences of the first protocol. They found no significant differences in sensitivity and specificity between each of the two abbreviated protocols (86 and 89%, respectively) and the full protocol (95%). However, the case series was relatively statistically underpowered while the enriched series (especially the proportion of malignant lesions, much higher than seen in screening practice) could have influenced the reader performance, likely explaining the remarkable low specificity, ranging from 45% to 52%.

In 2016, Laura Heacock and coworkers[54] retrospectively evaluated the utility of an abbreviated T1-weighted imaging protocol in detecting 107 known breast cancers (88% invasive and 12% in situ) as well as to analyze the impact of adding clinical history and prior imaging to cancer detection and determine the impact of T2-weighted imaging in cancer detection and lesion conspicuity. The abbreviated protocol, consisting of a T2-weighted fat-suppressed sequence and a pre- and post-contrast T1-weighted sequence, reached a sensitivity of 97.8 - 99.4%, comparable to previously mentioned studies[52,53,76]. In addition, in the Heacocks study[54], information about prior imaging and clinical history increased detection rates. T2-weighted imaging increased confidence and lesion conspicuity, however, it did not increase detection rates. Initial enhancement rate was significantly correlated with tumor grade, invasive disease, and lesion conspicuity, supporting the idea that rapid wash-in characteristics of malignancy may underpin the efficacy of abbreviated MRI sequences. This finding raises the possibility that cancers detected by an abbreviated MRI examination only may be of higher grade, i.e. more biologically active lesions, potentially counteracting the drawback of overdiagnosis intrinsically associated with every screening program.

From the studies investigating abbreviated protocols, we can conclude that there is still no clear consensus in which sequences are beneficial and needed for an ab-

Table 2.4: Abbreviated and full breast MRI protocols

First author	Protocol	Dynamic pre-contrast	Dynamic 1st post-contrast	Dynamic 2nd post-contrast	Dynamic 3rd-5th post-contrast	T1 TSE	T2	DWI
Kuhl[76]	Abbreviated	x	x					
	Full	x	x	x	x	x	x	
Mango[52]	Abbreviated	x	x					
	Full	x	x	x	x	x	x	x
Grimm[53]	Abbreviated1	x	x				x	
	Abbreviated2	x	x	x			x	
	Full	x	x	x	x	x	x	
Heacock[54]	Abbreviated	x	x				x	
	Full	x	x	x	x		x	x

breviated protocol. Studies reporting on abbreviated protocols varied widely in acquisition times ranging from approximately 3 minutes to 15 minutes. In particular, the role of a T2-weighted sequence for screening purposes needs to be investigated. In addition, the number of patients scanned and the number of cancers detected is currently still too low to draw solid conclusions. Future larger prospective trials need to prove the non-inferiority of abbreviated protocols. Nevertheless, shortened breast MRI protocols could play a relevant role in lowering costs and allowing more widespread availability of MRI as a screening tool.

2.8 Ultra-fast breast MRI

All the abbreviated protocols discussed in the previous paragraph discard dynamic information. Only one of those investigated by Grimm and coworkers[53] used the second dynamic post-contrast scan. This is not problematic for larger malignant lesions, which are generally well recognized based upon their morphological features. However, in particular for the classification of small mass lesions, which are typical findings in breast screening, additional dynamic information is important. This implies that dynamic information is appreciated, while imaging time should not be extended.

Conventional dynamic information cannot be obtained, as this requires acquisition of the wash-out phase of contrast which takes up to 6 - 7 minutes after contrast administration. However, even in the early days of breast MRI, it was already shown that dynamic information obtained from the inflow phase had better discriminating capacity than the wash-out phase. Nevertheless, in previous years the temporal resolution, typically in the range of 60 - 75 seconds of high spatial resolution bilateral images was not sufficient to document this inflow phase. Therefore, acquiring scans
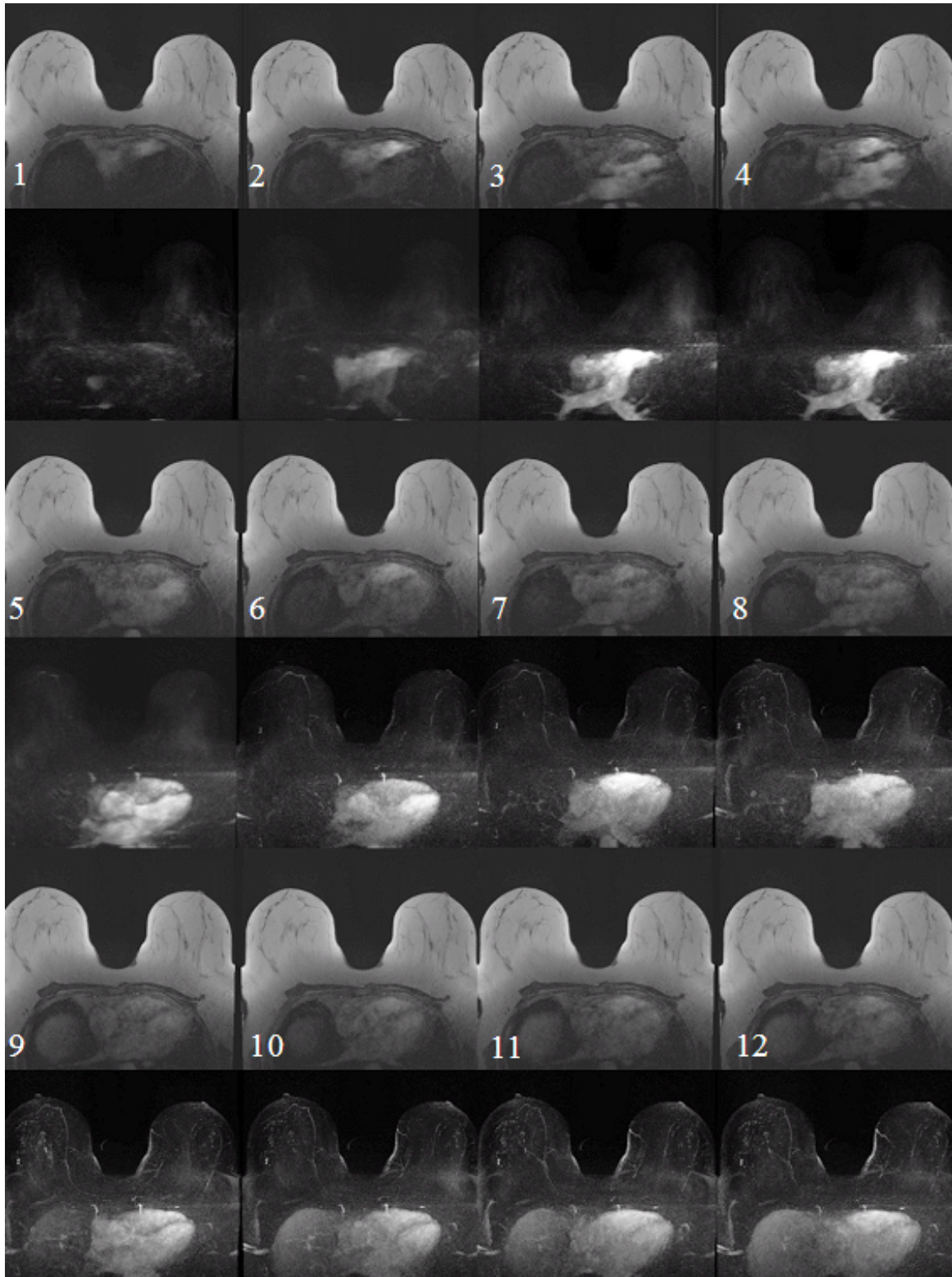
Figure 2.4: Ultrafast axial images in a 42-years-old *BRCA2* mutation carrier, using a TWIST sequence (temporal resolution 4.57s). The central slice of each of the first 12 acquisitions after aorta enhancement are shown, numbered from 1 to 12. The corresponding volumetric MIPs are shown at every time point, below the original central unsubtracted images. These clearly show the arrival of the contrast, first in the thoracic vessels and the heart, subsequently in the breasts and liver. No suspicious early enhancement is visible in the breasts.

at a high temporal resolution re-enables the use of contrast dynamics for the classification of suspicious breast lesions[77].

Karl-Heinz Herrmann and coworkers[78] were among the first to describe a new ultrafast sequence named time-resolved angiography with stochastic trajectories (TWIST) for breast MRI (see Figure 2.4). With this technique, the outer part of k-space is heavily under-sampled and data points are shared between successive time points to increase the obtained spatial resolution to diagnostic quality. Sophisticated sampling patterns are used to minimize the disadvantage of data sharing. These authors showed, in a pilot study of 14 patients, that this TWIST sequence can be used to obtain dynamic images at a very high temporal resolution (5.7 seconds). Furthermore, they showed that benign lesions enhance at a later time point than malignant lesions.

Luminita A. Tudorica and coworkers[79] reduced the temporal resolution to 18 s, showing that the dynamic images were very comparable with the images provided by the conventional protocol. Yuan Le and coworkers[80] showed that a TWIST sequence can be combined with a dual-echo (two-point) Dixon technique to obtain fat-suppressed images with a high temporal resolution. Our group[55] investigated the use of maximum slope of the contrast enhancement versus time curve obtained from the TWIST sequence at a temporal resolution of 4.3 seconds as a novel dynamic parameter for the differentiation between benign and malignant lesions. The total acquisition time was 102 seconds. Of the 199 enhancing lesions included, 95 were proven benign and 104 malignant. We found that maximum slope achieved a much higher accuracy in differentiating benign and malignant lesions than the BI-RADS curve type does, thus, solidifying the use of ultrafast breast MRI, and allowing the creation of new protocols with a short post-contrast period ($\sim$ 85 seconds). While we did not evaluate TWIST for morphological features, the technique meets every breast MRI requirement that is stated in guidelines. Results of a recently presented reader study in which four radiologists evaluated 200 screening cases showed that the use of ultrafast MRI alone was just as accurate as evaluating a full diagnostic protocol including high spatial resolution acquisitions, T2-weighted imaging, and DWI[81].

Federico D. Pineda and coworkers[77] investigated a bilateral, fat-suppressed ultrafast acquisition with a time resolution of 6.9-9.9 s during the first minute after contrast injection, followed by four high spatial resolution acquisitions with a time resolution of 60-79.5 s. They confirmed that first-minute ultrafast dynamic imaging can add valuable information, increasing the radiologists confidence in identifying lesions in the presence of marked background parenchymal enhancement. A hybrid

construction, where ultrafast acquisitions are interleaved in an abbreviated breast MRI protocol, allows the collection of dynamic data for lesion classification without a penalty in acquisition time.

Further improvements of ultrafast MRI are still increasing image quality. Radial imaging using a golden angle approach, as is performed in the golden-angle radial sparse parallel (GRASP) sequence, enables dynamic imaging using continuous data acquisition and retrospective reconstruction of image series with an arbitrary temporal resolution by grouping different numbers of consecutive radial lines into temporal frames. This means that, with the use of GRASP, images of every temporal resolution can be reconstructed, thus, both ultrafast and regular high spatial resolution acquisitions can be obtained using the same sequence, as described for liver, pediatrics, breast, and neck[82]. This approach can help to improve clinical workflow by enabling data acquisition without the need for synchronization with breath-hold commands or for selection of predefined rigid temporal resolution. A recently published study[83] showed that the performance of the GRASP sequence in terms of conspicuity of benign and malignant breast lesions is near-comparable to that of conventional volumetric imaging breath-hold examination (VIBE) imaging. Thus, techniques that employ compressed sensing might be used to further improve image quality of ultrafast imaging. Table 2.5 lists multiparametric protocols that include ultrafast MRI sequences recently described in literature.

2.9 Future perspectives: contrastless screening

As earlier discussed, one major disadvantage of breast MRI is the need for contrast agent administration. This is not solved by either abbreviated, ultrafast, or hybrid protocols. Since the recent observation of gadolinium deposition/retention in the brain in a fraction of patients who underwent multiple injections of gadolinium-based contrast agents[84–86], this has become a debated subject, especially for breast MRI screening, because healthy women at increased risk for breast cancer are annually exposed to gadolinium-based contrast agents.

Even though there are currently no clinical sequels associated with deposition/retention and guidelines for the use of breast MRI have remained unchanged[87], this is an additional reason (next to the associated costs, need for intravenous cannulation, and risks of already known adverse events such as allergic reactions) to investigate alternative MRI screening strategies that do not rely on contrast administration. The recent technological developments in cancer imaging have led to a shift towards

Table 2.5: Multiparametric protocols containing ultrafast sequences

First author[ref.#]	Protocol	
Pineda[77]	1.	Ultrafast sequence
	2.	High resolution sequences
Herrmann[78]	1.	Pre-contrast 2D gradient-echo
	2.	Ultrafast sequence (TWIST)
Tudorica[79]	1.	Ultrafast sequence (TWIST) with fat saturation
	2.	Axial T2-weighted sequence with fat saturation
	3.	3D T1-weighted sequence without fat saturation
Le[80]	Healthy volunteers, no contrast administration	
	1.	TWIST Dixon sequence
	2.	VIBE SPAIR
Mann[55]	1.	DWI
	2.	VIBE (T1-weighted)
	3.	TWIST (T1-weighted)
	4.	T2-weighted
Feng[82]	1.	GRASP
Heacock[83]	1.	Sagittal T1-weighted gradient echo
	2.	Sagittal T2-weighted gradient echo
	3.	Sagittal T1-weighted gradient echo
	When breast biopsy images were acquired; Sagittal GRASP sequence continuously acquired before, during, and after contrast injection	

functional assessment of tissue characteristics, possibly making contrast administration in the future unnecessary.

Of all techniques in use, DWI appears the strongest candidate for contrastless breast MRI. Like T2-weighted imaging, DWI does not require contrast agent administration, however, its sensitivity is much higher. This makes that DWI can be used in patients with a poor renal function or patients with an allergy to gadolinium-based contrast agents. Although currently not as good as contrast-enhanced MRI, the sensitivity of DWI is already competitive with that of mammography.

Sebastian Bickelhaupt and coworkers[88] investigated the fusion of T2-weighted images and DWI for characterization of BI-RADS 4/5 mammographic findings. Combining morphological information from the former with bio-physiological characteristics from the latter allowed radiologists to get a high diagnostic accuracy for lesion characterization (92%) comparable to that of the full DCE protocol (95%). Research is needed to investigate whether this can be used for breast MRI screening.

Further diffusion-based approaches to contrastless breast MRI include diffusion tensor imaging (DTI), that appears to improve the diagnostic capacity of DWI, and allows imaging at a substantially higher resolution than that is common for DWI[89] and intravoxel incoherent motion (IVIM), although the spatial resolution of the latter currently still precludes any screening[90].

An approach to obtain vascular information without contrast administration is arterial spin labeling (ASL). The arterial blood supplying the tissue of interest is labeled' altering its longitudinal magnetization; perfusion quantification can be easily performed as the signal changes are proportional to blood flow[91,92]. The technique has already been successfully implemented to improve disease detection and characterization in the brain, pancreas, and kidney[91]. Several ASL techniques exist, all based on three different spin states (equilibrium, saturation, and inversion) and can also be categorized as on- and off-slice tagging sequences. Pilot studies showed an ASL potential for distinguishing malignant from benign breast tissues (as malignant tumors had a higher water content than normal tissue and a higher perfusion than both normal tissues and benign lesions)[92] and a correlation between MRI perfusion values of breast tissue by ASL as compared to CT perfusion[93]. Unfortunately, the ASL sequences for breast are still under development and some cannot even cover the whole breast. In addition, the obtained signals are still too weak and reconstruction artefacts too strong to use it as a screening tool. It is a matter of time to see

whether ASL can evolve into a technique that can replace contrast-enhanced breast MRI.

Metabolic imaging, predominantly MRS, is another path for exploration. It is well established that total choline, in particular phosphocholine, is elevated in breast cancer[58,68–71,94]. As spectroscopic techniques improve, quantification of in-vivo spectra can be done more reliably. This replaces the criterion of seeing or not seeing the choline peak with more sophisticated quantitative criteria for judging whether a breast lesion might be malignant[58]. However, for future screening application, multivoxel techniques (2D, or better, 3D) are needed. Moreover, the SNR is so low that voxels are in the size order of cubic centimeters, thus, still not suitable for screening. Nevertheless, novel approaches to metabolic imaging, such as phosphorous spectroscopy and chemical saturation transfer imaging (CEST) may enable much higher resolution levels in the near future that might give them access to a new platform for screening research[95].

2.10 Conclusion

Breast MRI is solidly established as the most accurate screening technique for breast cancer available, even though currently mainly applied to women at increased risk of breast cancer. Current state-of-the-art protocols are multiparametric in nature and focus on achieving both a high sensitivity and a high specificity. As the added value of additional sequences on top of simple T1-weighted acquisitions in a screening setting appears questionable and is, likely, not cost-effective, current research focuses on shortening of MRI protocols. Both abbreviated and ultrafast approaches to breast MRI allow acquisition within minutes without losing in accuracy.

Future research focuses on the use of non-contrast techniques for screening. DWI currently seems most suitable. However, so far this technique cannot deliver the quality of screening obtained with contrast-enhanced techniques.

Screening performance of breast MRI and mammography

3

Suzan Vreemann, Albert Gubern-Merida, Margrethe S. Schlooz-Vries, Peter Bult, Carla H. van Gils, Nicoline Hoogerbrugge, Nico Karssemeijer, Ritse M. Mann

Abstract

Purpose To evaluate the real-life performance of a breast cancer screening program for women with different categories of increased breast cancer risk with multiple follow-up rounds in an academic hospital with a large screening population.

Materials and Methods Screening examinations (magnetic resonance [MR] imaging and mammography) for women at increased breast cancer risk (January 1, 2003, to January 1, 2014) were evaluated. Risk category, age, recall for workup of screening-detected abnormalities, biopsy, and histopathologic diagnosis were recorded. Recall rate, biopsy rate, positive predictive value of recall, positive predictive value of biopsy, cancer detection rate, sensitivity, and specificity were calculated for first and follow-up rounds.

Results There were 8818 MR and 6245 mammographic examinations performed in 2463 women. Documented were 170 cancers; of these, there were 129 screening-detected cancers, 16 interval cancers, and 25 cancers discovered at prophylactic mastectomy. Overall sensitivity was 75.9% including the cancers discovered at prophylactic mastectomy (95% confidence interval: 69.5%, 82.4%) and 90.0% excluding those cancers (95% confidence interval: 83.3%, 93.7%). Sensitivity was lowest for carriers of the *BRCA1* mutation (66.1% and 81.3% when including and not including cancers in prophylactic mastectomy specimens, respectively). Specificity was higher at follow-up (96.5%; 95% confidence interval: 96.0%, 96.9%) than in first rounds (85.1%; 95% confidence interval: 83.4%, 86.5%) and was high for both MR imaging (97.1%; 95% confidence interval: 96.7%, 97.5%) and mammography (98.7%; 95% confidence interval: 98.3%, 99.0%). Positive predictive value of recall and positive predictive value of biopsy were lowest in women who had only a family history of breast cancer.

Conclusion Screening performance was dependent on risk category. Sensitivity was lowest in carriers of the *BRCA1* mutation. The specificity of high-risk breast screening improved at follow-up rounds.

3.1 Introduction

Breast magnetic resonance imaging (MRI) is considered to be the most sensitive imaging modality for early breast cancer detection and is recommended as a supplemental screening technique for women with a lifetime risk for the development of breast cancer of 20 - 25% or higher[11,12]. This includes women with a *BRCA* germline mutation, for whom the lifetime risk is as high as 56 - 84%, and women who underwent radiation therapy to the chest wall at young age[11,12,37,96]. Furthermore, women with a strong family history for breast cancer, women with a personal history of breast cancer, and women with high-risk lesions such as atypical ductal hyperplasia and lobular carcinoma in situ are at increased risk, though for the latter risk categories the indication for additional screening with MRI is less clear[11,12,97]. Nonetheless, supplemental screening may be indicated for all these women because the rate of interval cancers is relatively high[98–100].

Whereas the sensitivity of breast MR imaging has been reported[101] to be as high as 95% in women with known breast cancer, the sensitivity of breast MRI in screening is lower. Initially, the sensitivity was reported to be around 77% in the screening setting, but more recent studies document a higher sensitivity of around 90%[15,41,75,102]. Studies[15,41,75] show a nearly doubled cancer detection of combined MR imaging and mammography screening compared with mammography screening alone. Consequently, high-risk screening programs with MR imaging and mammography have been implemented in clinical practices worldwide, although there are national, regional, and local differences regarding whom to screen with supplemental MR imaging[11,12,103]. This is likely also because it is so far unknown whether the diagnostic value of breast MR imaging screening is different for groups with a different underlying risk and whether the added value of MR imaging persists in follow-up evaluations for all groups.

A major criticism regarding breast MR imaging screening, beyond availability and cost, is that the specificity appears lower than generally accepted for mammography screening. Reported specificities range from 81% to 95% and compare poorly to a specificity of 99% for mammography screening[14,15,37,38,75,102,104]. However, to our knowledge, hardly any data exist on the effect of routine use of breast MR imaging and the availability of prior examinations on the diagnostic accuracy of breast MR imaging. Only a few recent studies of breast MR imaging screening programs document recall rates, false-positive findings, and interval cancers[15,41].

The purpose of this study is to evaluate the real-life performance of a breast cancer screening program for women with different categories of increased breast cancer risk with multiple follow-up rounds in an academic hospital with a large screening population.

3.2   Materials and Methods

Screening program

The breast cancer screening program for women at increased breast cancer risk ($\geq$ 20 - 25% lifetime risk) in carriers of the *BRCA* mutation. From the age of 30 or older, MR imaging is combined with a yearly mammographic examination in *BRCA* mutation carriers. In women at high familial risk, screening starts between 35 and 45 years or 5 years before the age at which the youngest relative developed breast cancer[20]. In women with a personal history of breast cancer, screening starts 1 year after the diagnosis of breast cancer.

Case selection

This retrospective study was approved by our local institutional review board and the requirement for informed consent was waived. The local database was searched to identify all screening MR examinations and mammographies performed from January 1, 2003 to January 1, 2014. Women were included when MR examinations were performed on the basis of a screening indication, mammograms were included if the indication was screening and mammography was performed within 1 year after the screening MR examination. We recorded risk category (*BRCA1*, *BRCA2*, family history, personal history, and other women), age, screening tests performed, recall for workup of screen-detected abnormalities, biopsies, and histopathologic diagnosis when available.

During the study period, 9571 screening breast MR examinations and 6553 screening mammographic examinations were performed in 2773 women. After excluding women not at increased risk, we evaluated 8818 screening breast MR examinations and 6245 screening mammographic examinations that were performed in 2463 women at increased risk for breast cancer (mean age, 44 $\pm$ 12 [standard deviation]; age range, 15 - 91). The population included 770 women who were *BRCA* mutation carriers (471 carriers of *BRCA1* and 299 carriers of *BRCA2*) and 26 untested first-degree relatives (16 carriers of *BRCA1* and 10 carriers of *BRCA2*). Furthermore, women were included on the basis of a family history of breast cancer (lifetime risk, > 20%; N = 748), a personal history of breast cancer (N = 836), or other reasons

Table 3.1: Characteristics of the study population

	BRCA1	*Untested first degree relatives BRCA1*	*BRCA2*	*Untested first degree relatives BRCA2*	*Family history*	*Personal history*	*Others at increased risk*
	(N=471)	**(N=16)**	**(N=299)**	**(N=10)**	**(N=748)**	**(N=836)**	**(N=83)**
Age at start mean (range)	39 (23-75)	31 (22-69)	41 (23-73)	29 (25-50)	42 (16-73)	52 (22-91)	38 (15-71)
Median no. of screening MRI exams performed (range)	3 (1-13)	2 (1-4)	3 (1-11)	1 (1-10)	2 (1-12)	2 (1-12)	3 (1-10)
Median no. of screening mammography exams performed (range)	2 (0-11)	1 (0-5)	3 (0-11)	1 (0-10)	1 (0-10)	1 (0-12)	1 (0-7)
VDG at start[$] (%)	398 available	11 available	265 available	7 available	526 available	474 available	52 available
a	88 (22.1)	1 (9.1)	52 (19.6)	0 (0.0)	46 (8.7)	59 (12.4)	10 (19.2)
b	104 (26.1)	1 (9.1)	63 (23.8)	2 (28.6)	83 (15.8)	130 (27.4)	19 (19.2)
c	117 (29.4)	4 (36.4)	102 (38.5)	1 (14.3)	202 (38.4)	187 (39.5)	21 (40.4)
d	89 (22.4)	5 (45.5)	48 (18.1)	4 (57.1)	195 (37.1)	98 (20.7)	11 (21.2)

[$] VDG: Volpara density grade. A volumetric and automatically assessed density score on the raw mammography data. The classes are created in analogy to the BI-RADS lexicon.

(chest radiation and high-risk lesions, N = 83). The mean number of MR examinations performed per woman was three (range, one to 13 examinations), and the mean number of mammographic examinations was two (range, zero to 12 examinations). Characteristics of the study population are presented in Table 3.1.

Image acquisition

MR imaging protocols varied over time. Examinations were performed on a 1.5- or 3-T imager (Magnetom Avanto, Magnetom Sonata, Magnetom Simphony, or Magnetom Trio; Siemens, Erlangen, Germany) by using a dedicated bilateral breast coil. Women were imaged in the prone position. A transverse or coronal three-dimensional T1-weighted gradient-recalled echo dynamic sequence was performed before contrast agent administration followed by four or five postcontrast-administration sequences. Pixel spacing, section thickness, matrix, echo time, repetition time, and flip angle differed among acquisitions as previously described by Dalmis et al.[105]. Diffusion-weighted imaging and T2-weighted acquisitions were added from those imaged in 2012. Various gadolinium chelates were used, administered at a dose of 0.1mmol/kg or 0.2mmol/kg by using a power-injector (Medrad, Warrendale, Pa) at a flow rate of 2.5mL/s, followed by a saline flush. Premenopausal women were

scheduled between the 6th and 12th day of their menstrual cycle.

Mammograms were obtained in two directions (mediolateral oblique and cranio-caudal) with a full-field digital mammography imager (GE Senograph 2000 or GE Senograph DS, GE, Fairfield, Conn). Additional views and spot-compression views were performed at request of the evaluating radiologist.

Image interpretation

All examinations were evaluated by one of 8 breast radiologists (R.M.M. and other radiologists) using the Breast Imaging Reporting and Data System (BI-RADS)[19,106]. Experience in breast MR imaging ranged from 0.5 to 23 years. For MR imaging, both morphologic and dynamic parameters were evaluated by using a dedicated breast MR imaging workstation (versions of DynaCAD, Invivo, Gainesville, Fla; and Philips, Best, the Netherlands). When available, mammograms, MR images, and previous examinations were evaluated together. Final assessment categories were given as specified in the American College of Radiology BI-RADS atlas[19]. Biopsies were performed for lesions classified as BI-RADS category 4 and BI-RADS category 5, and a subset of lesions classified as BI-RADS category 3. The remaining BI-RADS category 3 lesions underwent short-term follow-up.

Ground truth

Our database was linked to the nationwide population-based Netherlands Cancer Registry. Normal or benign screening examinations were confirmed by at least 1 year of clinical follow-up and regarded to be true-negative when no cancer was detected before the subsequent screening round. When no biopsy was indicated at short-term follow-up, at least one year of clinical follow-up was required to confirm benignity. All lesions that underwent biopsy and prophylactic mastectomies performed in-house were identified by a cross-computer search with our pathology database. We subsequently recorded whether the biopsy was performed for screening findings or symptoms.

Data-analysis

To evaluate the performance, we defined performance metrics (per examination) in analogy to those used for evaluation of mammography screening programs[33].

Screening examination negative for cancer were defined as screening examinations that were rated as BI-RADS category 1 or 2. True-negative screening examinations were defined as screening examinations negative for cancer in women in whom no

cancer was detected in the subsequent year.

Results of pathologic analysis were grouped into malignant (ie, ductal carcinoma in situ, invasive, and metastatic cancer) and benign (all other findings) lesions. Screen-detected cancers were defined as cancers diagnosed after diagnostic workup initiated by screening findings. In the high-risk screening program, we defined two groups of non-screening-detected cancers. The first group consisted of interval cancers that manifested in between screening rounds because of symptoms. The second group consisted of cancers that were detected in asymptomatic women who underwent prophylactic mastectomy. Prognostic differences between these groups were investigated using χ^2 tests for categorical variables or one-way analysis of variance for continuous variables.

Consequently, two different groups of false-negative examinations (referred to here as FN1 and FN2) were defined. FN1 consisted of screening examinations rated as BI-RADS category 1 or BI-RADS category 2 in which any cancer was detected in the following year or before the subsequent screening examination. FN2 consisted of screening examinations rated as BI-RADS category 1 or BI-RADS category 2 in which a true interval cancer was detected.

The recall rate was defined as the number of examinations that were classified as BI-RADS category 0, BI-RADS category 3, BI-RADS category 4, or BI-RADS category 5 per 1000 screening examinations performed. The biopsy rate was defined as the number of examinations in which a biopsy was performed per 1000 screening examinations performed. The cancer detection rate was defined as the number of examinations that led to breast cancer detection per 1000 screening examinations.

Screening examinations positive for cancer were defined as screening examinations that were rated as BI-RADS category 0, BI-RADS category 3, BI-RADS category 4, or BI-RADS category 5, and true-positive examinations were defined as screening examinations positive for cancer that led to cancer detection. False-positive examinations were defined as examinations that led to a recall in women in whom no breast cancer was detected. The positive predictive value of recall was defined as the fraction of recalls that led to cancer detection. The positive predictive value of biopsy was defined as the fraction of biopsies that led to cancer detection.

We calculated the sensitivity of the complete screening program and for both modalities separately. Three different sensitivities were recognized: the sensitivity for all

breast cancers detected, the sensitivity for screening-detected cancers and interval cancers detected because of symptoms (excluding cancers detected in prophylactic mastectomies), and the modality-specific sensitivity for screening-detected cancers. For the latter category, lesions were considered visible with a modality when they were mentioned in radiologic report of that modality (or report section when both modalities were reported simultaneously). Cancers only reported in mammography were considered to be false-negative findings in MR imaging and vice versa. Specificity was determined using examinations of women that had at least 1 year of follow-up without recall and calculated for complete screening and independently for each modality.

We analyzed the overall screening results and separated results of the first screening round from follow-up rounds. The first screening round was defined as the first time a woman with increased breast cancer risk underwent MR imaging. Follow-up screening rounds were defined as screening examinations performed in a period between 10 and 24 months after a previous screening round.

To explore the influence of risk category on screening performance, the dataset was divided into risk subgroups. To compare first screening rounds with follow-up rounds, and analyze differences between subgroups, Fisher exact tests and one-way analysis of variance were used and 95% confidence intervals were calculated[107]. Additionally, the difference in tumor stage, nodal stage, and grade between cancers detected in the first round and cancers detected in follow-up rounds were assessed using χ^2 statistics. Bonferroni correction was applied for performance measures, and a two-sided P-value of ≤ 0.008 or less was considered to indicate statistical significance. All statistics were performed by using statistical software (SPSS version 22; SPSS, Chicago, Ill).

3.3 Results

A total of 577 women (mean age, 44 $\pm$ 11; age range, 15 - 77 years) were recalled, some more than once (505 women were recalled once, 60 women were recalled twice, nine women were recalled three times, one woman was recalled four times, and two women were recalled five times), which resulted in 666 recalls. This led to 424 women undergoing 475 biopsies (one biopsy, 383 women; two biopsies, 35 women; three biopsies, three women; four biopsies, two women; five biopsies, one woman; mean age, 44 years $\pm$ 11; age range, 15 - 77 years) as a consequence of screening.

Cancers

In total, 170 newly diagnosed cancers in 156 women were documented. Of these, 129 (75.9%) cancers were screening-detected and 118 cancers were detected at MR imaging (modality sensitivity, 91.5%; 47 cancers were detected at the first round and 71 cancers were detected at follow-up rounds). Mammography was available for 108 screening-detected cancers, and in 62 women a cancer was visible at mammography (modality sensitivity, 57.4%). Eleven of these cancers were found at mammography alone (all at follow-up; mean age, 52 years $\pm$ 8.8; age range, 35 - 69 years), of which seven cancers (63.6%) were pure ductal carcinoma in situ. The incremental cancer detection at mammography was stronger (although not significant) in women without a *BRCA* mutation than in those with a *BRCA* mutation (sensitivity increase in *BRCA* mutation carriers vs in women without a *BRCA* mutation who underwent both examinations, 3.3% [from 64.1% to 67.4%] vs 10.3% [from 75.6% to 85.9%], respectively; P = 0.11).

Forty-one interval cancers were detected outside screening, of which 16 were found because of patient symptoms. Twenty-five cancers were found at prophylactic mastectomy. Of these cancers, 21 were detected at our hospital among 421 prophylactic mastectomies in 246 patients (cancer frequency per breast, 5.0%). The frequency of prophylactic mastectomy varied among risk categories, and most were performed in carriers of the *BRCA1* mutation (265 mastectomies in *BRCA1*, 111 mastectomies in *BRCA2*, 16 mastectomies in women with a family history, 27 mastectomies in women with a personal history, and two mastectomies in the group with other reasons). Eighteen of 25 (72.0%) cancers detected at prophylactic mastectomy were pure ductal carcinoma in situ. Patient and tumor characteristics of screening-detected and interval cancers are shown in Table 3.2. In general, symptomatic interval cancers have a higher tumor stage, are of higher grade, and are more often invasive compared with those found at prophylactic mastectomy (P $\leq$ 0.05).

Screening performance

Performance metrics are presented in Table 3.3 with individual risk category in Tables 3.5 - 3.11. The overall screening program had a sensitivity of 75.9% when all recognized cancers were included and a sensitivity of 90.0% when prophylactic mastectomy cancers were excluded. Whereas the sensitivity of MR imaging for invasive disease was higher than that of mammography (78.7% vs 46.2%, respectively, with inclusion of all cancers [P < 0.001]; 83.3% vs 48.0%, respectively, excluding cancers in prophylactic mastectomies [P < 0.001]), the sensitivities of mammography and MR imaging for ductal carcinoma in situ were not significantly different (45.2% vs

41.9%, respectively, with the inclusion of all cancers [P = 0.82], and 70.0% vs 72.0%, respectively, excluding cancers detected in prophylactic mastectomies [P = 0.99]; see Table 3.4).

The sensitivity of the screening program appears to be higher in carriers of the *BRCA2* mutation (69.7% including cancers found at prophylactic mastectomy and 92.0% excluding cancers found at prophylactic mastectomy), women with a family history positive for cancer (90.9% and 95.2% with and without cancers found at prophylactic mastectomy, respectively), women with a personal history of breast cancer (82.0% and 91.1% with and without cancers found at prophylactic mastectomy, respectively), and other patients at an increased level (100% both with and without cancers found at prophylactic mastectomy) compared with carriers of the *BRCA1* mutation (66.1% and 81.3% with and without cancers found at prophylactic mastectomy, respectively), though this did not reach statistical significance (P = 0.05 and P = 0.25 with and without cancers found at prophylactic mastectomy, respectively).

Overall recall ranged from 5.8 to 10.3% (from 57.8/1000 to 103.2/1000) across subpopulations and was substantially lower in follow-up examinations than in first round examinations (P < 0.001, see Table 3.3). Positive predictive value of recall was 0.19 and was lowest for women with a family history of breast cancer (positive predictive value of recall, 0.09) (see Table 3.9). At follow-up rounds, the overall cancer detection rate was lower in all subgroups, which shows the large fraction of prevalent cancers detected at the first screening round. The largest decrease in cancer detection rate was found in carriers of the *BRCA1* mutation and other at increased breast cancer risk (a difference of 23.0 [P = 0.008] and 42.1 [P = 0.05], respectively). Nonetheless, we did not detect a difference in pT stage (P = 0.12), pN stage (P = 0.63), or tumor grade (P = 0.33) between cancers detected at the first round and follow-up rounds. Despite the declining cancer detection rate in follow-up rounds because of a lower prevalence of cancer, sensitivity of MR imaging (69.4%) was higher than that of mammography (45.9%; P < 0.001). Recall rate and biopsy rate strongly declined in follow-up rounds, whereas positive predictive value of recall and positive predictive value of biopsy increased for the entire population at both MR imaging (P = 0.004 and P = 0.005, respectively) and mammography (P < 0.001 for both positive predictive values) except for the subgroup of other patients at increased breast cancer risk (see Table 3.11). For MR imaging, the availability of previous imaging examinations (MR imaging and/or mammography) strongly improved the specificity to a level that was just below the specificity of mammography (97.1% vs 98.7%; P < 0.001).

3.4 Discussion

We showed that the specificity of high-risk screening in real-life practice is excellent, both for mammography and breast MR imaging, especially in follow-up rounds. This is different than in previous clinical trials that reported variable and generally much lower specificity of breast MR imaging[14,15,37,102]. In follow-up rounds, the specificity of MR imaging approached the specificity of mammography, which reaffirmed that new examinations must be evaluated together with previous studies. By doing so, the number of recalls is reduced and the positive predictive value of both recall and biopsy increases.

The positive predictive value is substantially lower in women with a family history of breast cancer than in all other subgroups. This might be partly related to the much lower breast cancer risk in this population. Our results thus still indicate that strict evaluation of risk category for intensified screening is required.

In terms of sensitivity, excluding cancers detected in prophylactic mastectomy, the overall performance of the screening regimen is 90.0% and varied little between subgroups, except for carriers of the *BRCA1* mutation in whom the sensitivity (81.3%) appeared to be lower. This is in line with earlier studies[39] and implies that, for carriers of the *BRCA1* mutation, an even more stringent screening regimen, or prophylactic mastectomy[108], may be indicated.

Despite the high accuracy of the screening program, it is evident that there is a reservoir of undetected cancers that are only found in prophylactic mastectomy specimens. At our center, 5% of prophylactic mastectomies contained undetected cancers. Hence, the overall sensitivity of the screening program is heavily dependent on the frequency of prophylactic mastectomies in the target population and the indication for prophylactic mastectomy. However, our results showed that the cancers found in prophylactic mastectomy specimens were in general small and usually non-invasive, which is in line with findings of previous studies[109,110]. Consequently, the importance of detecting these cancers is unknown, even though they are most often detected in carriers of the *BRCA* germline mutation, which, in turn, can be partly explained by the fact that more prophylactic surgical procedures are performed in *BRCA* mutation carriers.

The addition of mammography to MR imaging increased the sensitivity of the screening program to 8.1% (from 71.9% with only MR imaging to 80.0% with both exam-

inations) in women who underwent both examinations. These were 11 cancers, of which seven were ductal carcinoma in situ. Whether the additional detection of these cancers is beneficial to the patient is unknown. Kuhl et al[18] noted that the sensitivity of breast MR imaging for high-grade ductal carcinoma in situ is excellent, whereas mammography preferentially depicts lower grade lesions[18]. In line with other studies, our results show that the sensitivity of mammography was especially low in carriers of the *BRCA* mutation and in women with a personal history of breast cancer[102,111–113]. Nevertheless, the age from which supplemental mammography screening is recommended varies from country to country and may be as low as 30 years in carriers of the *BRCA* mutation[11]. This advice could be modified to reflect the importance of MR screening in this population.

Our study has limitations. This was a single institutional study, which potentially limits its ability to generalize. Because the study is longitudinal in nature, clinical and imaging protocols evolved over time. Furthermore, 13 radiologists reported the cases over the years. Unfortunately, because of the limited number of cancer detections per year, the influence of changing protocols and different levels of experience cannot be assessed. Whereas we analyzed the performance of MR imaging and mammography separately, in most women these examinations were simultaneously evaluated. Therefore, the performance of each of the modalities may have been affected by the findings in the other modality. In addition, because we only have information on the prophylactic mastectomies negative for cancer that were performed at our hospital, the actual frequency of cancer in specimens from prophylactic mastectomy might be slightly over- or understated because of selection bias. Another limitation is the fact that young women who are *BRCA* mutation carriers (< 30 years) are only screened with MR imaging, and hence the number of MR examinations was different from the number of mammographic examinations, which could potentially skew accuracy values. Finally, the rate of prophylactic mastectomy reduces the apparent sensitivity for all cancers, so any difference in sensitivity in breast cancer patients may largely be because of the higher rate of such mastectomies rather than an unique difficulty in detection.

From our results it is evident that the group of women who underwent intensified screening because of a family history of breast cancer was at substantially lower risk than the other subgroups. Because of the retrospective nature of the study, we were unable to obtain the actual family history and hence could not further stratify these women. Similarly, the subgroup with risk category "other" remains heterogeneous, but because of the low frequency of the specific risks within this group, further divi-

sion was not possible. Finally, survival data were not available. Further studies need to establish whether there is a survival benefit for women who participate in breast cancer screening programs[114,115].

In conclusion, this study shows that, in real-life practice, the performance of a high-risk screening program is affected by risk category and the frequency of prophylactic mastectomies. The specificity of the screening program is high and improves at follow-up. Especially in women who are *BRCA* mutation carriers, the added value of mammography is limited.

Table 3.2: Patient and cancer characteristics of cancers detected in high-risk screening patients

	Screen-detected cancers	Interval cancers due to patient symptoms	Cancers found at prophylactic mastectomy	P-value
Number of cancers (%)	129 (75.9)	16 (9.4)	25 (14.7)	N/A
Median age at detection in years (range in years)	50 (24-77)	42 (27-55)	42 (27-70)	<0.001*
Risk category				0.10
- BRCA1 mutation (%)	39 (66.1)	9 (15.3)	11 (18.6)	
- Untested *BRCA1** (%)	0 (0.0)	0 (0.0)	0 (0.0)	
- BRCA2 mutation (%)	23 (69.7)	2 (6.1)	8 (24.2)	
- Untested *BRCA2** (%)	0 (0.0)	0 (0.0)	0 (0.0)	
- Family history (%)	20 (90.9)	1 (4.5)	1 (4.5)	
- Personal history (%)	41 (82.0)	8 (8.0)	5 (10.0)	
- Others (%)	6 (100.0)	0 (0.0)	0 (0.0)	
Invasive disease	104 (81.9)	16 (12.6)	7 (5.5)	N/A
Mean pathological size$^{\#}$ (range in mm)	14.2 (2-70)	15.5 (5-26)	5.8 (1-11)	0.24
T-stage				0.54
- pT1 (%)	78 (83.9)	8 (8.6)	7 (7.5)	
- pT2 (%)	18 (85.7)	3 (14.3)	0 (0.0)	
- pT3 (%)	1 (100.0)	0 (0.0)	0 (0.0)	
- pT4D (%)	1 (100.0)	0 (0.0)	0 (0.0)	
- Recurrence (%)	4 (66.7)	2 (33.3)	0 (0.0)	
- Unknown (%)	2 (40.0)	3 (60.0)	0 (0.0)	
Grade				0.17
- Grade 1 (%)	19 (90.5)	0 (0.0)	2 (9.5)	
- Grade 2 (%)	31 (86.1)	4 (11.1)	1 (2.8)	
- Grade 3 (%)	45 (80.4)	8 (14.3)	3 (5.4)	
- Unknown (%)	9 (64.3)	4 (28.6)	1 (7.1)	
Molecular subtype				0.01
- Luminal A (%)	32 (88.9)	3 (8.3)	1 (2.8)	
- Luminal B (%)	20 (95.2)	0 (0.0)	1 (4.8)	
- HER2-type (%)	9 (100.0)	0 (0.0)	0 (0.0)	
- Triple negative (%)	18 (60.0)	8 (26.7)	4 (13.3)	
- Unknown (%)	25 (80.6)	5 (16.1)	1 (3.2)	
N-stage				0.15
- pN0 (%)	71 (80.7)	7 (8.0)	4 (4.5)	
- pN+ (%)	24 (82.8)	5 (17.2)	0 (0.0)	
- Unknown (%)	9 (56.3)	4 (25.0)	3 (18.8)	
M-stage				N/A
- pM0 (%)	52 (82.5)	5 (7.9)	6 (9.5)	
- Unknown (%)	52 (81.3)	11 (17.2)	1 (1.6)	
In situ disease (%)	25 (58.1)	0 (0.0)	18 (41.9)	N/A
Grade				0.02
- Grade 1 (%)	1 (20.0)	0 (0.0)	4 (80.0)	
- Grade 2 (%)	12 (52.2)	0 (0.0)	11 (47.8)	
- Grade 3 (%)	9 (90.0)	0 (0.0)	1 (10.0)	
- Unknown (%)	3 (60.0)	0 (0.0)	2 (40.0)	
N-stage				0.40
- pN0 (%)	21 (61.8)	0 (0.0)	13 (38.2)	
- pN+ (%)	0 (0.0)	0 (0.0)	1 (100.0)	
- Unknown (%)	4 (50.0)	0 (0.0)	4 (50.0)	

* Cancers detected in untested first degree relatives of *BRCA* mutation carriers

$^{\#}$ Pathological tumor size is known for 88 screen-detected invasive lesions, for 8 interval cancers, and for 5 cancers detected in prophylactic mastectomy specimen.

In situ disease; pure ductal carcinoma in situ

Table 3.3: Performance measures for mammography and MRI screening in the entire population

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=8818	N=2092	N=6726	
RR (per 1000)	75.5 [71.5-82.2]	167.8 [152.6-183.0]	46.8 [42.1-52.0]	<0.001*
BR (per 1000)	53.9 [50.2-59.4]	113.3 [100.2-126.0]	35.4 [31.6-40.3]	<0.001*
PPV1	0.19 [0.16-0.22]	0.13 [0.10-0.16]	0.26 [0.21-0.31]	<0.001*
PPV3	0.27 [0.23-0.30]	0.20 [0.14-0.24]	0.34 [0.28-0.40]	<0.001*
CDR (per 1000)	14.6 [12.1-17.0]	22.5 [15.8-27.9]	12.2 [9.7-15.0]	0.001*
Sensitivity1 (%)	75.9 [69.5-82.4]	81.0 [69.6-90.2]	73.2 [64.7-81.2]	0.35
Sensitivity2 (%)	90.0 [83.3-93.7]	88.7 [77.4-95.6]	89.1 [81.4-94.7]	0.99
Specificity (%)	93.8 [93.3-94.3]	85.1 [83.4-86.5]	96.5 [96.0-96.9]	<0.001*
Mammography	N=6245	N=1645	N=4600	
RR (per 1000)	38.1 [36.3-46.0]	82.1 [74.5-101.1]	22.4 [19.4-28.2]	<0.001*
BR (per 1000)	24.7 [22.9-30.8]	46.2 [39.5-60.0]	17.0 [14.6-22.4]	<0.001*
PPV1	0.26 [0.20-0.31]	0.14 [0.09-0.20]	0.42 [0.33-0.52]	<0.001*
PPV3	0.40 [0.32-0.47]	0.25 [0.16-0.34]	0.55 [0.44-0.65]	<0.001*
CDR (per 1000)	9.9 [8.1-13.2]	11.6 [7.4-18.0]	9.3 [7.4-13.3]	0.47
Sensitivity1 (%)	45.9 [35.5-51.4]	41.3 [27.3-54.9]	48.3 [35.2-54.8]	0.47
Sensitvity 2 (%)	51.7 [43.0-60.7]	45.2 [30.5-59.8]	55.1 [44.7-66.6]	0.34
Specificity (%)	97.1 [96.7-97.5]	92.7 [91.4-93.9]	98.7 [98.3-99.0]	<0.001*
MRI	N=8818	N=2092	N=6726	
RR (per 1000)	61.6 [57.9-67.7]	134.3 [120.2-148.0]	39.0 [34.9-44.0]	<0.001*
BR (per 1000)	48.1 [44.3-53.0]	103.7 [90.2-115.0]	30.8 [27.0-35.2]	<0.001*
PPV1	0.22 [0.18-0.24]	0.17 [0.12-0.20]	0.27 [0.22-0.32]	0.004*
PPV3	0.28 [0.23-0.31]	0.22 [0.15-0.26]	0.34 [0.28-0.40]	0.005*
CDR (per 1000)	13.4 [10.8-15.5]	22.5 [15.5-27.4]	10.6 [8.3-13.1]	<0.001*
Sensitivity1 (%)	69.4 [61.9-75.8]	81.0 [67.8-89.0]	63.4 [54.1-72.1]	0.02
Sensitivity2 (%)	81.4 [73.7-86.7]	88.7 [75.3-94.4]	77.2 [67.5-85.0]	0.12
Specificity (%)	95.1 [94.6-95.5]	88.5 [87.0-89.8]	97.1 [96.7-97.5]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.4: Sensitivity for invasive disease and DCIS for mammography and MRI

	Mammography	MRI	P-value
Invasive disease			
- Sensitivity1	46.2 [36.8-55.7] (48/104)	78.7 [70.8-85.0] (100/127)	<0.001*
- Sensitivity2	48.0 [38.5-57.7] (48/100)	83.3 [75.7-88.9] (100/120)	<0.001*
DCIS*			
- Sensitivity1	45.2 [29.2-562.2] (14/31)	41.9 [28.4-56.7] (18/43)	<0.001*
- Sensitivity2	70.0 [48.1-85.5] (14/20)	72.0 [52.4-85.7] (18/25)	<0.001*

* DCIS; ductal carcinoma in situ

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomy specimen as false negatives

Table 3.5: Performance measures for mammography and MRI screening in *BRCA1 mutation carriers (excluding untested first degree relatives)*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=1714	N=424	N=1290	
RR (per 1000)	76.4 [64.8-89.9])	155.7 [124.3-193.3]	50.4 [39.7-63.7]	<0.001*
BR (per 1000)	58.3 [48.2-70.4])	115.6 [88.6-149.5]	39.5 [30.2-51.6]	<0.001*
PPV1	0.30 [0.23-0.38]	0.26 [0.17-0.37]	0.34 [0.24-0.46]	0.34
PPV3	0.39 [0.30-0.49]	0.35 [0.23-0.49]	0.43 [0.31-0.57]	0.42
CDR (per 1000)	22.8 [16.7-31.0]	40.1 [25.2-63.3]	17.1 [11.3-25.7]	0.008*
Sensitivity1 (%)	66.1 [53.4-76.9]	73.9 [53.5-87.5]	61.1 [44.9-75.2]	0.40
Sensitivity2 (%)	81.3 [68.1-89.8]	77.3 [56.6-89.9]	84.6 [66.5-93.9]	0.71
Specificity (%)	94.4 [93.2-95.4]	87.8 [84.2-90.6]	96.6 [95.4-97.4]	<0.001*
Mammography	N=1534	N=395	N=1139	
RR (per 1000)	29.3 [22.0-39.0]	53.2 [35.1-79.9]	21.1 [14.2-31.2]	0.10
BR (per 1000)	21.5 [15.3-30.0]	32.9 [19.3-55.5]	17.6 [11.4-27.0]	0.11
PPV1	0.49 [0.35-0.63]	0.33 [0.17-0.55]	0.63 [0.43-0.79]	0.08
PPV3	0.67 [0.50-0.80]	0.54 [0.29-0.77]	0.75 [0.53-0.89]	0.27
CDR (per 1000)	14.3 [9.5-21.6]	17.7 [8.6-36.1]	13.2 [8.0-21.7]	0.47
Sensitivity1 (%)	44.9 [31.9-58.7]	36.8 [19.2-59.0]	50.0 [33.2-66.9]	0.40
Sensitvity 2 (%)	51.2 [36.8-65.4]	38.9 [20.3-61.4]	60.0 [40.7-76.6]	0.22
Specificity (%)	98.4 [97.7-99.0]	96.3 [93.8-97.8]	99.2 [98.5-99.6]	<0.001*
MRI	N=1714	N=424	N=1290	
RR (per 1000)	66.5 [55.6-79.3]	136.8 [107.3-136.8]	43.4 [33.6-55.9]	<0.001*
BR (per 1000)	53.1 [43.4-64.7]	108.5 [82.3-141.7]	34.9 [26.2-46.4]	<0.001*
PPV1	0.32 [0.25-0.42]	0.29 [0.19-0.42]	0.36 [0.24-0.49]	0.55
PPV3	0.41 [0.31-0.51]	0.37 [0.25-0.51]	0.44 [0.31-0.59]	0.53
CDR (per 1000)	21.6 [15.7-29.6]	40.1 [25.2-63.3]	15.5 [10.1-23.8]	0.006*
Sensitivity1 (%)	62.7 [50.0-73.9]	73.9 [53.5-87.5]	55.6 [39.6-70.5]	0.18
Sensitivity2 (%)	77.1 [63.5-86.7]	77.3 [56.6-89.9]	76.9 [58.0-89.0]	0.99
Specificity (%)	95.3 [94.2-96.3]	89.8 [86.4-92.4]	97.1 [96.1-97.9]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.6: Performance measures for mammography and MRI screening in *BRCA1 mutation carriers (including untested first degree relatives)*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=1748	N=440	N=1308	
RR (per 1000)	77.2 [65.6-90.7]	159.1 [127.9-196.2]	49.7 [39.2-62.9]	<0.001*
BR (per 1000)	58.9 [48.8-70.9]	118.2 [91.3-151.7]	39.0 [29.8-50.9]	<0.001*
PPV1	0.29 [0.22-0.37]	0.24 [0.16-0.36]	0.34 [0.24-0.46]	0.26
PPV3	0.38 [0.29-0.48]	0.33 [0.22-0.46]	0.43 [0.31-0.57]	0.31
CDR (per 1000)	22.3 [16.4-30.3]	38.6 [24.2-61.0]	16.8 [11.1-25.3]	0.01
Sensitivity1 (%)	66.1 [53.4-76.9]	73.9 [53.5-87.5]	61.1 [44.9-75.2]	0.40
Sensitivity2 (%)	81.3 [68.1-89.8]	77.3 [56.6-89.9]	84.6 [66.5-93.9]	0.71
Specificity (%)	94.3 [93.1-95.3]	87.3 [83.8-90.2]	96.6 [95.5-97.5]	<0.001*
Mammography	N=1552	N=405	N=1147	
RR (per 1000)	29.6 [22.3-39.3]	54.3 [36.1-80.8]	20.9 [14.1-30.9]	0.002*
BR (per 1000)	21.3 [15.2-29.8]	32.1 [18.9-54.1]	17.4 [11.3-26.7]	0.11
PPV1	0.48 [0.34-0.62]	0.32 [0.16-0.53]	0.63 [0.43-0.79]	0.05
PPV3	0.67 [0.50-0.80]	0.54 [0.29-0.77]	0.75 [0.53-0.89]	0.27
CDR (per 1000)	14.2 [9.4-21.4]	17.3 [8.4-35.3]	13.1 [7.8-21.5]	0.62
Sensitivity1 (%)	44.9 [31.9-58.7]	36.8 [19.2-59.0]	50.0 [33.2-66.9]	0.40
Sensitvity 2 (%)	51.2 [36.8-65.4]	38.9 [20.3-61.4]	60.0 [40.7-76.6]	0.22
Specificity (%)	98.4 [97.6-98.9]	96.1 [93.7-97.6]	99.2 [98.5-99.6]	<0.001*
MRI	N=1748	N=440	N=1308	
RR (per 1000)	66.9 [56.1-79.6]	138.6 [109.4-174.0]	42.8 [33.1-55.2]	<0.001*
BR (per 1000)	53.8 [44.2-65.4]	111.4 [85.3-144.2]	34.4 [25.8-45.7]	<0.001*
PPV1	0.32 [0.24-0.41]	0.28 [0.18-0.40]	0.36 [0.24-0.49]	0.43
PPV3	0.39 [0.30-0.49]	0.35 [0.23-0.49]	0.44 [0.31-0.59]	0.40
CDR (per 1000)	21.2 [15.4-29.1]	38.6 [24.2-61.0]	15.3 [9.9-23.5]	0.006*
Sensitivity1 (%)	62.7 [50.0-73.9]	73.9 [53.5-87.5]	55.6 [39.6-70.5]	0.18
Sensitivity2 (%)	77.1 [63.5-86.7]	77.3 [56.6-89.9]	76.9 [58.0-89.0]	0.99
Specificity (%)	95.3 [94.1-96.2]	89.4 [96.1-92.1]	97.2 [96.1-98.0]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.7: Performance measures for mammography and MRI screening in *BRCA2 mutation carriers (excluding untested first degree relatives)*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=1123	N=278	N=845	
RR (per 1000)	83.7 [68.9-101.4]	190.6 [148.8-240.9]	48.5 [35.9-65.1]	<0.001*
BR (per 1000)	67.7 [54.4-83.9]	140.3 [104.4-186.0]	43.8 [31.9-59.8]	<0.001*
PPV1	0.24 [0.17-0.34]	0.13 [0.07-0.25]	0.39 [0.26-0.54]	0.007*
PPV3	0.30 [0.21-0.41]	0.18 [0.09-0.33]	0.43 [0.29-0.59]	0.02
CDR (per 1000)	20.5 [13.7-30.6]	25.2 [12.3-51.1]	18.9 [11.7-30.5]	0.48
Sensitivity1 (%)	69.7 [52.7-82.6]	77.8 [45.3-93.7]	66.7 [46.7-82.0]	0.69
Sensitivity2 (%)	92.0 [75.0-97.8]	100.0 [64.6-100.0]	88.9 [67.2-96.9]	0.99
Specificity (%)	93.5 [91.9-94.8]	82.9 [78.0-86.9]	97.0 [95.0-97.9]	<0.001*
Mammography	N=1009	N=258	N=751	
RR (per 1000)	33.7 [24.2-46.7]	65.9 [41.5-103.0]	22.6 [14.2-35.9]	0.002*
BR (per 1000)	26.8 [18.5-38.7]	46.5 [26.8-79.5]	21.3 [13.2-34.2]	0.05
PPV1	0.32 [0.19-0.49]	0.18 [0.06-0.41]	0.47 [0.26-0.69]	0.14
PPV3	0.41 [0.25-0.59]	0.27 [0.10-0.57]	0.50 [0.28-0.72]	0.43
CDR (per 1000)	10.9 [6.1-19.4]	11.6 [3.9-33.6]	10.7 [5.4-20.9]	0.99
Sensitivity1 (%)	35.5 [21.1-53.1]	33.3 [12.1-64.6]	36.4 [19.7-57.0]	0.99
Sensitivty 2 (%)	44.0 [26.7-62.9]	42.9 [15.8-75.0]	44.4 [24.6-66.3]	0.99
Specificity (%)	97.6 [96.5-98.4]	94.4 [90.8-96.6]	98.8 [97.7-99.4]	<0.001*
MRI	N=1123	N=278	N=845	
RR (per 1000)	74.8 [60.8-91.7]	172.7 [132.8-221.5]	42.6 [30.9-58.4]	<0.001*
BR (per 1000)	61.4 [48.8-77.0]	129.5 [95.0-174.1]	39.1 [28.0-54.4]	<0.001*
PPV1	0.26 [0.18-0.36]	0.15 [0.07-0.27]	0.42 [0.27-0.58]	0.007*
PPV3	0.32 [0.22-0.44]	0.19 [0.10-0.35]	0.45 [0.30-0.62]	0.04
CDR (per 1000)	19.6 [13.0-29.5]	25.2 [12.3-51.1]	17.8 [10.8-29.1]	0.46
Sensitivity1 (%)	66.7 [49.6-80.3]	77.8 [45.3-93.7]	62.5 [42.7-78.8]	0.68
Sensitivity2 (%)	88.0 [70.0-95.8]	100.0 [64.6-100.0]	83.3 [60.8-94.2]	0.53
Specificity (%)	94.3 [92.8-95.5]	84.8 [80.0-88.6]	97.4 [96.1-98.3]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.8: Performance measures for mammography and MRI screening in *BRCA2 mutation carriers (including untested first degree relatives)*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=1154	N=288	N=866	
RR (per 1000)	85.8 [71.0-103.4]	194.4 [152.8-244.0]	49.7 [37.1-66.3]	<0.001*
BR (per 1000)	67.6 [54.5-83.6]	138.9 [103.7-183.6]	43.9 [32.2-59.7]	<0.001*
PPV1	0.23 [0.16-0.32]	0.13 [0.06-0.24]	0.37 [0.24-0.52]	0.007*
PPV3	0.29 [0.21-0.40]	0.18 [0.09-0.32]	0.42 [0.28-0.58]	0.03
CDR (per 1000)	19.9 [13.3-29.7]	24.3 [11.8-49.3]	18.5 [11.4-29.8]	0.63
Sensitivity1 (%)	69.7 [51.7-82.6]	77.8 [45.3-93.7]	66.7 [46.7-82.0]	0.69
Sensitivity2 (%)	92.0 [75.0-97.8]	100.0 [64.6-100.0]	88.9 [67.1-96.9]	0.99
Specificity (%)	93.2 [91.6-94.5]	82.4 [77.5-86.5]	96.8 [95.4-97.8]	<0.001*
Mammography	N=1031	N=265	N=766	
RR (per 1000)	33.0 [23.7-45.8]	64.2 [40.5-100.4]	22.2 [13.9-35.3]	0.002*
BR (per 1000)	26.2 [18.1-37.8]	41.5 [23.3-72.8]	20.9 [12.9-33.7]	0.12
PPV1	0.32 [0.19-0.49]	0.18 [0.06-0.41]	0.47 [0.26-0.69]	0.14
PPV3	0.41 [0.25-0.59]	0.27 [0.10-0.57]	0.50 [0.28-0.72]	0.43
CDR (per 1000)	10.7 [6.0-19.0]	11.3 [3.8-32.7]	10.4 [5.3-20.4]	0.99
Sensitivity1 (%)	35.5 [21.1-53.1]	33.3 [12.1-64.6]	36.4 [19.7-57.0]	0.99
Sensitvity 2 (%)	44.0 [26.7-62.9]	42.9 [15.8-75.0]	44.4 [24.6-66.3]	0.99
Specificity (%)	97.7 [96.6-98.5]	94.5 [91.0-96.7]	98.8 [97.7-99.4]	<0.001*
MRI	N=1154	N=288	N=866	
RR (per 1000)	77.1 [63.1-93.9]	177.1 [137.3-255.4]	43.9 [32.2-59.7]	<0.001*
BR (per 1000)	61.5 [49.0-76.9]	128.5 [94.7-172.1]	39.3 [28.3-54.4]	<0.001*
PPV1	0.25 [0.17-0.35]	0.14 [0.07-0.26]	0.39 [0.21-0.55]	0.007*
PPV3	0.31 [0.21-0.42]	0.19 [0.09-0.34]	0.44 [0.29-0.61]	0.04
CDR (per 1000)	19.1 [12.7-28.7]	24.3 [11.8-49.3]	17.3 [10.5-28.4]	0.46
Sensitivity1 (%)	66.7 [49.6-80.3]	77.8 [45.3-93.7]	62.5 [42.7-78.8]	0.68
Sensitivity2 (%)	88.0 [70.0-95.8]	100.0 [64.6-100.0]	83.3 [60.8-94.2]	0.53
Specificity (%)	94.0 [92.5-95.3]	84.2 [79.5-88.0]	97.3 [95.9-98.2]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.9: Performance measures for mammography and MRI screening in *women with a family history of breast cancer*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=2624	N=613	N=2011	
RR (per 1000)	87.3 [76.9-98.9]	189.2 [159.4-223.0]	56.2 [46.7-67.4]	<0.001*
BR (per 1000)	55.6 [47.3-65.3]	109.3 [86.3-137.4]	39.3 [30.2-47.1]	<0.001
PPV1	0.09 [0.06-0.13]	0.05 [0.02-0.11]	0.12 [0.07-0.20]	0.06
PPV3	0.14 [0.09-0.21]	0.09 [0.04-0.19]	0.18 [0.10-0.28]	0.15
CDR (per 1000)	7.6 [4.8-12.0]	9.8 [4.0-22.3]	7.0 [4.0-12.0]	0.44
Sensitivity1 (%)	90.9 [69.4-98.4]	100.0 [51.7-100.0]	87.5 [60.4-97.8]	0.99
Sensitivity2 (%)	95.2 [74.1-99.8]	100.0 [51.7-100.0]	93.3 [66.0-99.7]	0.99
Specificity (%)	92.0 [90.9-93.0]	81.9 [78.6-84.7]	95.0 [94.0-95.9]	<0.001*
Mammography	N=1885	N=467	N=1418	
RR (per 1000)	47.7 [38.8-58.6]	113.5 [86.9-146.6]	26.1 [18.7-36.2]	<0.001*
BR (per 1000)	22.8 [16.8-30.9]	42.8 [27.0-66.5]	16.2 [10.6-24.6]	0.002*
PPV1	0.13 [0.07-0.23]	0.04 [0.01-0.14]	0.27 [0.14-0.44]	0.003*
PPV3	0.28 [0.16-0.44]	0.10 [0.02-0.33]	0.43 [0.24-0.65]	0.02
CDR (per 1000)	6.4 [3.5-11.4]	4.3 [0.7-17.1]	7.1 [3.6-13.4]	0.74
Sensitivity1 (%)	63.2 [38.6-82.8]	50.0 [9.0-90.8]	66.7 [38.7-87.0]	0.60
Sensitvity 2 (%)	66.7 [41.2-85.6]	50.0 [9.2-90.8]	71.4 [42.0-90.4]	0.56
Specificity (%)	95.8 [94.8-96.6]	90.0 [85.8-91.5]	98.1 [97.2-98.7]	<0.001*
MRI	N=2624	N=613	N=2011	
RR (per 1000)	64.0 [55.1-74.2]	128.9 [103.9-158.6]	44.3 [35.9-54.4]	<0.001*
BR (per 1000)	49.2 [41.4-58.3]	101.1 [79.0-128.4]	33.3 [26.1-42.4]	<0.001*
PPV1	0.10 [0.06-0.15]	0.08 [0.03-0.16]	0.11 [0.06-0.20]	0.45
PPV3	0.12 [0.07-0.20]	0.10 [0.04-0.21]	0.15 [0.08-0.26]	0.43
CDR (per 1000)	6.1 [3.6-10.1]	9.8 [4.0-22.3]	5.0 [2.5-9.5]	0.23
Sensitivity1 (%)	72.7 [49.6-88.4]	100.0 [51.7-100.0]	62.5 [35.9-83.7]	0.13
Sensitivity2 (%)	76.2 [52.5-90.9]	100.0 [51.7-100.0]	66.7 [38.7-87.0]	0.26
Specificity (%)	94.2 [93.2-95.0]	88.0 [85.1-90.3]	96.0 [95.1-96.8]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.10: Performance measures for mammography and MRI screening in *women with a personal history of breast cancer*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=3011	N=674	N=2337	
RR (per 1000)	57.8 [49.8-66.9]	135.0 [110.6-163.7]	35.5 [28.6-44.0]	<0.001*
BR (per 1000)	40.5 [33.9-48.3]	90.5 [70.4-115.4]	26.1 [20.2-33.6]	<0.001*
PPV1	0.24 [0.18-0.31]	0.14 [0.08-0.24]	0.34 [0.24-0.45]	<0.001*
PPV3	0.34 [0.25-0.43]	0.21 [0.12-0.34]	0.46 [0.33-0.59]	0.001*
CDR (per 1000)	13.6 [9.9-18.6]	19.3 [10.8-33.6]	12.0 [8.1-17.5]	0.18
Sensitivity1 (%)	82.0 [68.1-91.0]	81.3 [53.7-95.0]	82.4 [64.8-92.6]	0.99
Sensitivity2 (%)	91.1 [77.9-97.1]	92.9 [64.1-99.6]	90.3 [73.1-97.5]	0.99
Specificity (%)	95.5 [94.7-96.2]	88.1 [85.5-90.4]	97.6 [96.9-98.2]	<0.001*
Mammography	N=1658	N=458	N=1200	
RR (per 1000)	35.6 [27.4-46.0]	78.6 [56.4-108.1]	19.2 [12.5-29.1]	<0.001*
BR (per 1000)	25.9 [19.1-35.1]	56.8 [38.1-83.1]	14.2 [8.5-23.1]	<0.001*
PPV1	0.24 [0.14-0.37]	0.14 [0.05-0.30]	0.39 [0.20-0.61]	0.03
PPV3	0.33 [0.20-0.49]	0.19 [0.07-0.40]	0.53 [0.29-0.76]	0.04
CDR (per 1000)	8.4 [4.8-14.5]	10.9 [4.0-26.8]	7.5 [3.7-14.7]	0.55
Sensitivity1 (%)	40.0 [24.4-57.8]	45.5 [18.1-75.4]	37.5 [19.6-59.2]	0.72
Sensitvity 2 (%)	46.7 [28.8-65.4]	55.6 [22.7-84.7]	42.9 [22.6-65.6]	0.70
Specificity (%)	97.2 [96.3-97.9]	93.1 [90.4-95.1]	98.8 [98.0-99.3]	<0.001*
MRI	N=3011	N=674	N=2337	
RR (per 1000)	47.2 [40.0-55.5]	108.3 [86.4-134.8]	29.5 [23.2-37.4]	<0.001*
BR (per 1000)	34.2 [28.1-41.5]	75.7 [57.4-99.0]	22.3 [16.8-29.3]	<0.001*
PPV1	0.26 [0.19-0.34]	0.16 [0.09-0.27]	0.36 [0.25-0.49]	0.008*
PPV3	0.36 [0.27-0.46]	0.24 [0.13-0.38]	0.48 [0.34-0.62]	0.01
CDR (per 1000)	12.3 [8.8-17.1]	17.8 [9.7-31.8]	10.7 [7.1-16.0]	0.16
Sensitivity1 (%)	74.0 [59.4-84.9]	75.0 [47.4-91.7]	73.5 [55.3-86.5]	0.99
Sensitivity2 (%)	82.2 [67.4-91.5]	85.7 [56.2-97.5]	80.7 [61.9-91.9]	0.99
Specificity (%)	96.5 [95.7-97.1]	90.7 [88.3-92.7]	98.1 [97.5-98.6]	<0.001*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

Table 3.11: Performance measures for mammography and MRI screening in *others at increased risk*

	Overall [95% CI]	First screening round [95% CI]	Follow-up rounds [95% CI]	P-value (comparison first round vs. follow-up rounds)
Complete regimen[1]	N=281	N=77	N=204	
RR (per 1000)	103.2 [72.8-144.3]	233.8 [153.3-339.6]	53.9 [30.4-93.9]	<0.001*
BR (per 1000)	92.5 [36.9-132.1]	220.8 [142.7-325.5]	44.1 [23.4-81.7]	<0.001*
PPV1	0.21 [0.10-0.38]	0.22 [0.09-0.45]	0.18 [0.05-0.48]	0.99
PPV3	0.23 [0.11-0.42]	0.24 [0.10-0.47]	0.22 [0.06-0.55]	0.99
CDR (per 1000)	21.4 [9.9-45.9]	51.9 [20.4-126.0]	9.8 [2.7-35.0]	0.05
Sensitivity1 (%)	100.0 [61.0-100.0]	100.0 [51.0-100.0]	100.0 [34.2-100.0]	N/A
Sensitivity2 (%)	100.0 [61.0-100.0]	100.0 [51.0-100.0]	100.0 [34.2-100.0]	N/A
Specificity (%)	91.6 [87.8-94.4]	80.8 [70.3-88.2]	95.5 [91.7-97.6]	<0.001*
Mammography	N=119	N=50	N=69	
RR (per 1000)	75.6 [40.3-137.5]	140.0 [69.5-261.9]	29.0 [8.0-99.7]	0.03
BR (per 1000)	67.2 [34.4-127.0]	120.0 [56.2-238.0]	29.0 [8.0-99.7]	0.07
PPV1	0.33 [0.12-0.65]	0.29 [0.08-0.64]	0.50 [0.09-0.91]	0.99
PPV3	0.38 [0.14-0.69]	0.33 [0.10-0.70]	0.50 [0.09-0.91]	0.99
CDR (per 1000)	25.2 [8.6-71.5]	40.0 [11.0-134.6]	14.5 [2.6-77.6]	0.57
Sensitivity1 (%)	60.0 [23.1-88.2]	50.0 [15.0-85.0]	100.0 [20.7-100.0]	0.99
Sensitvity 2 (%)	60.0 [23.1-88.2]	50.0 [15.0-85.0]	100.0 [20.7-100.0]	0.99
Specificity (%)	94.7 [89.0-79.6]	89.1 [77.0-95.3]	98.5 [92.1-99.7]	0.04
MRI	N=281	N=77	N=204	
RR (per 1000)	96.1 [66.9-136.2]	207.8 [132.2-311.2]	53.9 [30.4-93.9]	<0.001*
BR (per 1000)	89.0 [61.0-128.1]	207.8 [132.2-311.2]	44.1 [23.4-81.7]	<0.001*
PPV1	0.22 [0.11-0.41]	0.25 [0.10-0.50]	0.18 [0.05-0.48]	0.99
PPV3	0.24 [0.12-0.43]	0.25 [0.10-0.50]	0.22 [0.06-0.55]	0.99
CDR (per 1000)	21.4 [9.9-45.9]	51.9 [20.4-126.0]	9.8 [2.7-35.0]	0.05
Sensitivity1 (%)	100.0 [61.0-100.0]	100.0 [51.0-100.0]	100.0 [34.2-100.0]	N/A
Sensitivity2 (%)	100.0 [61.0-100.0]	100.0 [51.0-100.0]	100.0 [34.2-100.0]	N/A
Specificity (%)	92.4 [88.6-95.0]	83.6 [73.4-90.3]	95.5 [91.7-97.6]	0.003*

[1] The complete regimen consist of a combination of MRI and mammography when available

RR: Recall Rate; BR: Biopsy Rate; PPV1: Positive Predictive Value1; PPV3: Positive Predictive Value3; CDR: Cancer Detection Rate

Sensitivity1: Sensitivity including all cancers

Sensitivity2: Sensitivity excluding cancers detected in prophylactic mastectomies as false negatives

Specificity: Specificity including cancers detected in prophylactic mastectomies as false negatives

The 95% confidence intervals (95% CI) are presented between brackets behind each of the presented performance measures.

The added value of mammography when breast MRI is available

Suzan Vreemann, Jan C.M. van Zelst, Margrethe M. Schooz-Vries, Peter Bult, Nicoline Hoogerbrugge, Nico Karssemeijer, Albert Gubern-Merida, Ritse M. Mann

4

Abstract

Introduction Breast MRI is the most sensitive imaging method for breast cancer detection and is therefore offered as a screening technique to women at increased risk of developing breast cancer. However, mammography is currently still added from the age of 30 without proven benefits. The purpose of this study is to investigate the added value of mammography when breast MRI is available.

Materials and Methods This retrospective single-center study evaluated 6553 screening rounds for 2026 women at increased risk (01/01/2003 - 01/01/2014). Risk category (*BRCA* mutation versus others at increased risk of breast cancer), age at exam, recall, biopsy, and histopathological diagnosis were recorded. Cancer yield, false positive recall rate (FPR), and false positive biopsy rate (FPB) were calculated using Generalized Estimating Equations for age-categories (< 40, 40 - 50, 50 - 60, ≥ 60 years) separately. Numbers of screens needed to detect an additional breast cancer with mammography (NSN) were calculated for each of the subgroups.

Results Of a total of 125 screen-detected breast cancers, 112 were detected by MRI and 66 by mammography. 13 cancers were solely detected by mammography, including 8 ductal carcinoma in situ-cases, most in women ≥ 50 years (77%). In higher age groups, mammography added more to the FPR. Below 50 years the mean number of mammographic examinations needed to find an MR-occult cancer was 1427.

Conclusion Mammography is of limited added value when breast MRI is available for women at increased risk of all ages. While in older women the additional cancer detection increases, there is also an increase in false positive findings.

4.1   Introduction

Mammography based screening for breast cancer reduces breast cancer-related mortality in the general female population[116]. However, in women at increased risk (e.g. those with a germline mutation in the *BRCA1* or *BRCA2* genes) biennial mammographic screening is insufficient due to low sensitivity and high rates of interval cancers[14,117–119]. Consequently, these women at a higher-than-average lifetime breast cancer risk (approximately $\geq$ 20 - 25% lifetime risk (LTR)) are invited to intensified screening programs[11,12], consisting of dynamic contrast-enhanced MRI (DCE-MRI) and mammography. The sensitivity and specificity of these screening programs have been reported to be as high as 97% and 98%, respectively[14,15,37,39,75,102].

Recent studies question the added value of mammography in this population, especially in *BRCA* mutation carriers[113]. In the study of Kuhl et al.[75], MRI proved to be the most important contributor to stage reduction. Although these results show the superiority of breast MRI compared to mammography for the detection of cancers, routine mammography is currently recommended for all women, even at a relatively young age. Various authors have proposed to cancel mammographic screening in young women also screened with breast MRI, especially in *BRCA1* mutation carriers. In these *BRCA1* mutation carriers, the mammographic sensitivity is exceedingly low, reported as low as 35%[120]. This is believed to be caused not only by the on average dense breasts of these women, but also by the mammographic benign-like features of *BRCA1*-associated cancers[121,122]. Berrington de Gonzalez et al. reported that there is little to no benefit of mammographic screening under the age of 35[123]. Additionally, concerns are raised for the risk of radiation-induced cancers in these women, since an increased radiation susceptibility exists in *BRCA* mutation carriers[123,124].

Although guidelines may vary per country, mammographic screening in *BRCA* mutation carriers is already advised from an age of 30 years[11,12]. However, the actual benefits in terms of tumor detection of the addition of mammography at such a young age are still unclear. In addition, additional findings on the mammogram might lead to an increase in false-positive recalls of the screening program.

Hence, there is a clinical need to find an optimal regimen for intensified screening programs to prevent unnecessary recalls, biopsies, and radiation exposure. The purpose of this study is to evaluate the added value of mammography when breast MRI is available in terms of cancer detection and false positive rates in a population of

women at increased risk of developing breast cancer.

4.2   Materials and Methods

This retrospective study was approved by our local institutional review board and the requirement for informed consent was waived.

Screening program

The increased risk screening program was evaluated for the period 01/01/2003 until 01/01/2014. The program starts at age 25 years for *BRCA* mutation carriers with yearly MRI. At the age of 30, a yearly mammography is added. Women with an LTR of $\geq$ 20 - 25% are from the start screened with mammography and MRI, starting ages differ by the reason for screening[20]. Furthermore, women may have been enrolled in the program at a later point in time, after detection of a specific factor that increases their personal risk. We previously reported on the overall screening performance in this cohort[125].

Case selection

The local database was searched to identify all screening MRI and mammography examinations. Women were included when an MRI examination was considered a screening examination (inquiry at the radiology department was for screening purposes in asymptomatic women). Women were excluded when no mammography was performed within six months of the screening MRI. Risk category, age, screening tests performed, eventual recall for workup of screen-detected abnormalities and histopathological diagnosis were recorded when available.

Image Acquisition

MRI acquisitions and protocols varied over time and were reported in detail in a previous report[105]. In short: examinations were performed on either a 1.5 or 3.0 Tesla Siemens scanner (Magnetom Avanto, Magnetom Sonata, Magnetom symphony or Magnetom Trio) using a dedicated bilateral breast coil. Patients were imaged in prone position. A transverse or coronal three-dimensional T1-weighted gradient-echo dynamic sequence was performed before contrast agent administration followed by 4 or 5 post-contrast sequences. Various gadolinium chelates were used as a contrast agent, administered at a dose of 0.1 mmol/kg or 0.2 mmol/kg using a power injector (Medrad, Warrendale, PA) at a flow rate of 2.5 mL/s, followed by a saline flush. Premenopausal women were scheduled in the 6 - 12th day of their menstrual cycle.

Mammograms were obtained in two directions (medio-lateral oblique and cranio-caudal) with a full field digital mammography machine (GE Senograph 2000 or GE Senograph DS, GE, Fairfield, Connecticut, USA). Additional views and spot compression views were performed at the request of the evaluating radiologist.

Image Interpretation

The Breast Imaging Reporting and Data system (BI-RADS)[19,126] was used for evaluation. All exams were evaluated by one of eight breast radiologists with experience ranging from 0.5 - 23 years after certification. Reporting was done using a dedicated breast MRI workstation (versions of DynaCAD, Invivo, Philips, Best, the Netherlands). Mammograms were evaluated together with MRI examinations when these examinations were acquired the same day. In general, biopsies were performed for lesions classified as BI-RADS 4 and 5, and a subset of lesions classified as BI-RADS 3. The remainder of BI-RADS 3 lesions underwent short-term follow-up.

Ground truth

For BI-RADS 3 lesions with short-term follow-up recommendation, at least one year of clinical follow-up was required to confirm benignity. A cross-computer search of our pathology records was performed to identify all performed biopsies. We subsequently analyzed if the biopsy was triggered by screening findings or whether the woman presented with symptoms. To ensure detection of all cancers, the database was also linked to the nationwide population-based Netherlands Cancer Registry (NCR).

Data-analysis

Pathology results were grouped into malignant (in situ, invasive and metastatic cancer) and benign lesions (all other findings). Only screen-detected cancers were investigated, which were defined as cancers diagnosed after diagnostic workup initiated by screening findings. We separated screen-detected cancers by mammography, MRI, or both based on radiological reports of the respective modalities (or report sections when mammogram and MRI were reported simultaneously).

Cancer yield, false positive recall rate (FPR) and false positive biopsy rate (FPB) for mammography, MRI, and the combination were calculated. Cancer yield was defined as the number of screen-detected cancers per 1000 screening rounds. An FPR or FPB was defined as a recalled/biopsied woman who was considered disease-free after work-up and/or after at least one year of clinical follow-up. The FPR/ FPB

were defined as the number of FPRs/FPBs per 1000 screening rounds.

Two risk categories were evaluated (*BRCA* mutation carriers, and all others). The *BRCA* mutation carriers group also included first-degree, untested relatives. Examinations were grouped into four age categories to investigate the influence of age ($<$ 40, 40 - 50, 50 - 60, $\geq$ 60 years).

Statistical analysis

Descriptive statistics were extracted. χ^2 tests were applied to compare differences between groups in proportion of ductal carcinoma in situ (DCIS) and tumor grade. To investigate the distribution of parameters over age categories χ^2 trend-tests were performed. Repeated screening results were summarized to form binomial counts for each woman to estimate cancer yield, FPR, and FPB. For each woman, the number of true positive and true negative screens per modality, and the number of screening visits with or without breast cancer detected were counted. In this way, binomial counts per modality were calculated and analyzed. As the dependent variable was assumed to follow a binomial distribution, Generalized Estimating Equations (GEE) was applied. The binomial proportions were modeled and conducted separately for cancer yield, FPR, and FPB, using a compound symmetry correlation structure. The analysis was conducted separately for each age category, modality, and risk category. After applying Bonferroni correction, a two-sided P-value of 0.013 was considered statistically significant. The number of mammography screens needed (NSN) to detect one breast cancer that was missed by MRI was calculated by dividing the number of mammography screens performed by the number of breast cancers detected by mammography alone. All statistics were performed by SPSS (version 22, SPSS Inc., Chicago, IL).

4.3   Results

Study population

Final analysis included 2026 women with 6553 screening rounds (see Table 4.1). 125 screen-detected cancers were identified of which 13 and 59 were only detected by mammography or MRI, respectively (P < 0.001). In total, 112 cancers were seen on MRI and 66 on mammography. Overall, no significant difference was found between tumor grade of cancers detected by mammography or MRI (P = 0.193). Mammography detected a significantly higher proportion of pure DCIS (16/66 (24%) and 15/112 (13%) for mammography or MRI, respectively, P < 0.001). We did not observe a difference in grade of DCIS detected with mammography or MRI (P = 0.436).

Mammography detected breast cancers
The majority of cancers detected only with mammography consisted of pure DCIS (pTis) (8/13, 62%, see Table 4.2). Most women who were diagnosed with pure DCIS were ≥ 50 years of age (6/8, 75%, see Table 4.2). The remaining five women with

Table 4.1: Population and breast cancer characteristics in the cohort

	<40 years	40 - 50 years	50 - 60 years	≥ 60 years	Overall	P-value[*]
Women (N)						
- BRCA	388	258	182	75	903	<0.001
- Others	329	504	482	273	1588	0.014
- Overall	717	762	664	348	2491	<0.001
Exams (N)						
- BRCA	1113	737	568	190	2608	<0.001
- Others	716	1313	1265	651	3945	0.046
- Overall	1829	2050	1833	841	6553	<0.001
BC (N)						
- Mammography	13	13	25	15	66	0.253
- Mammography only	1	2	8	2	13	0.202
- MRI	25	30	37	20	112	0.697
- MRI only	13	19	20	7	59	0.254
- Overall	26	32	45	22	125	0.963
Invasive tumor (N)						
- Mammography	13	9	16	12	50	0.771
- Mammography only	1	0	4	0	5	0.822
- MRI	24	25	30	18	97	0.496
- MRI only	12	16	18	6	52	0.253
- Overall	25	25	34	18	102	0.540
DCIS (N)						
- Mammography	0	4	9	3	16	0.073
- Mammography only	0	2	5	1	8	0.281
- MRI	1	5	7	2	15	0.036
- MRI only	1	3	2	1	7	0.848
- Overall	1	7	11	4	23	0.164
Tumor grade of all cancers (N)						
Grade 1						
- Mammography	1	1	4	1	7	0.536
- Mammography only	0	0	1	0	1	0.655
- MRI	3	4	5	5	17	0.384
- MRI only	2	3	2	4	11	0.442
Grade 2						
- Mammography	1	5	7	4	17	0.171
- Mammography only	0	1	2	1	4	0.317
- MRI	3	10	17	6	36	0.170
- MRI only	2	6	12	3	23	0.335
Grade 3						
- Mammography	11	7	11	6	35	0.339
- Mammography only	1	1	2	1	5	0.822
- MRI	17	15	15	5	52	0.010
- MRI only	7	9	6	0	22	0.009
Missing	2	1	3	4	10	0.197
FPR (N)						
- Mammography	63	75	55	22	215	<0.001
- Mammography only	22	38	28	15	103	0.115
- MRI	159	143	72	25	399	<0.001
- MRI only	118	106	45	18	287	<0.001
- Overall	181	181	100	40	502	<0.001
FPB (N)						
- Mammography	35	46	32	9	122	<0.001
- Mammography only	6	15	11	3	35	0.258
- MRI	114	113	51	18	296	<0.001
- MRI only	85	82	30	12	209	<0.001
- Overall	120	128	62	21	331	<0.001

BC = breast cancer (invasive cancer and ductal carcinoma in situ (DCIS)); FPR = false positive recall; FPB = false positive biopsy

[*] χ^2 test for trend was performed for the fraction of the overall population

an invasive cancer detected only at mammography were aged 35, 53, 54, 55 and 56. All pure mammography-detected breast cancers were detected in follow-up rounds. The NSN for the overall population and the defined subgroups are presented in Table 4.3. There was no cancer that was not reported by MRI in the first rounds of screening, making an estimate of NSN not applicable. Our results show that the NSN was highest in the lowest age categories. Whether a difference exists between women with a proven *BRCA* mutation and women without is difficult to make, since we did not observe only mammographically detected breast cancers in *BRCA* mutation carriers under 50 years of age.

Cancer yield

Cancer yield increased over time, with a peak at the 50 - 60 age category (see Figure 4.1). The difference between cancers detected by MRI and the combination (mammography + MRI) seemed to increase with age ($<$ 40 years: 0.47, 40 - 50 years: 0.93, 50 - 60 years: 4.26, $\geq$ 60 years: 2.93 per 1000 examinations), pointing to a possible increased added value of mammography in higher age categories (see Figure 4.2), which was the strongest in the 50 - 60 years categories for both *BRCA* mutation carriers and others. The increase in breast cancer yield by the addition of mammography was not significant in all risk categories ($P \geq 0.303$). Tables 4.4 - 4.6 summarize cancer yield, FPR, and FPB.

False positives

For FPRs, mammography added 103 FPRs on top of 112 FPRs based on both mammography and MRI, and 287 FPRs based on MRI alone. Overall, mammography significantly added to the FPRs ($P = 0.001$), especially in the group of women without a *BRCA* mutation ($P = 0.001$). The relative increase of the FPR due to mammography was higher in higher age-groups ($<$ 40 years: 14%, 40 - 50 years: 27%, 50 - 60 years: 44%, $\geq$ 60 years: 61%, see Figure 4.2). This was significant in women without a *BRCA* mutation ($P < 0.001$). In total 35 FPBs were performed based on mammography alone. This did not lead to a significant increase in the overall FPBs ($P = 0.013$), or in any of the subcategories ($P \geq 0.323$).

Completely omitting mammography from the screening regimen would have led to a reduction of 21% (103/502) of FPRs and 11% (35/331) of FPBs.

Table 4.2: Breast cancers detected solely by mammography

No.	Risk category	Ipsi/Contra$	Age	Tumor type	Tumor size*	Tumor grade	ER-status	PR-status	H2N-status	Nodal status	1st round versus FU
1	BRCA1	N/A	50	DCIS	6	2	-	-	-	0	FU
2	Family	N/A	43	DCIS	7	2	-	-	-	0	FU
3	Family	N/A	48	DCIS	-	3	-	-	-	0	FU
4	Personal	Ipsi	55	DCIS	23	-	-	-	-	0	FU
5	Personal	Contra	58	DCIS	-	-	-	-	-	0	FU
6	Personal	Contra	69	DCIS	6	2	-	-	-	0	FU
7	Other2	N/A	61	DCIS	21	3	-	-	-	-	FU
8	Personal	Contra	55	DCIS	-	3	-	-	-	0	FU
9	BRCA2	N/A	57	DCIS [1]	6	-	Positive	Positive	Negative	0	FU
10	BRCA1	N/A	56	IDC	8	3	Positive	Positive	Negative	0	FU
11	Family	N/A	35	IDC	4	3	Positive	Positive	-	0	FU
12	Family	N/A	53	Tubular	3	1	Positive	Positive	Negative	0	FU
13	Other2	N/A	54	ILC	23	2	Positive	Positive	Negative	1mi	FU

- = not available; N/A = not applicable

* = pathological tumor size (in mm), in case of multi-centric tumors (case 13) the diameter of the largest tumor is mentioned

$ Breast cancer in the ipsi- or contralateral breast in patients with a personal history of breast cancer

[1] DCIS with micro-invasive growth

4.4   Discussion

This study evaluated the added value of mammography on top of MRI in a multi-modal imaging screening program for women who are at intermediate or high risk of developing breast cancer in a single academic institute. The addition of mammography translated mostly to the detection of a small number of DCIS cases that were occult on MRI. However, five additional invasive carcinomas were also detected. The number of mammography screening exams needed to detect an MRI occult cancer depended on age and was very high in women under 40. In addition, adding mammography led to a slight increase in false-positive recalls and biopsies.

Screening, with the aim of early detection of (pre-) malignant breast lesions to decrease breast cancer-related mortality, is a well-accepted risk-reducing strategy for most women at increased risk of developing breast cancer[127]. MRI is considered the most accurate imaging modality[15,32,41,75]. Mammography is currently added to most screening regimens that include MRI to detect calcified breast lesions that may be visualized with mammography but not with MRI[128,129]. In our study, eight out of 13 cancers (62%) were MRI occult ductal in situ cancers that were detected based on microcalcifications on the mammogram. The five invasive cancers that were detected only with mammography, were also found because of microcalcifications. By mammography alone, only 1 invasive cancer was detected in a *BRCA* mutation carrier, at the age of 56. Our results are in line with the meta-analysis of Heijnsdijk et al.[130], who reported only one invasive cancer detected by mammography alone in *BRCA1* mutation carriers across 4 breast cancer screening trials of women at high risk for developing breast cancer. Obdeijn et al.[113] also reported little benefit of mammography screening in younger women with a *BRCA1* mutation. In their study, omitting mammography from the screening regimen would have led to two missed DCIS cases in women aged 50 and 67. Obdeijn et al. suggested to increase the starting age for mammography screening in women with *BRCA1* mutations to 40 years.

Our results suggest that the detection of MRI occult breast cancers is very rare in all women younger than 40 years. Ten out of 13 MRI occult cancers (both DCIS and invasive cancers) were observed in women ≥ 50 years in our population, which is in line with the results reported by Narayan et al.[131]. In our study, raising the starting age of mammography to 40 years would have led to missing one invasive ductal cancer in a woman with a positive family history of breast cancer but without a known *BRCA* mutation and no DCIS cases would have been missed. In retrospect, this invasive cancer was visible on the MRI and was therefore not truly occult accord-

ing to Gubern-Merida et al.[132]. Additionally, while in older women the additional detection of breast cancer increases with the addition of mammography, this is counterbalanced by an increase in false positive findings. These results are supported by the data of Phi and coworkers[111].

Other imaging modalities may be used to detect additional cancers on top of MRI. Unfortunately, handheld ultrasound or even automated breast ultrasound has been shown to be of limited value in a screening setting where MRI is available[15,32,75,102,133]. Digital breast tomosynthesis (DBT) has also been shown to increase the cancer detection rate and decrease the number of FPRs when compared to mammography alone in women at average risk[25,134]. However, there is no consensus on the added value of DBT when breast MRI is available[135]. Therefore, current guidelines only include mammography. The gain in sensitivity with mammography seems to come mostly from the detection of lesions presenting with calcifications. DBT appears to be of relatively equal value to mammography for this purpose, but at a higher dose[27,30]. Since younger women at high risk and in particular *BRCA* mutation carriers have been shown to be more susceptible to develop radiation-induced cancers[123,124], replacing mammography with DBT might not be beneficial for women screened with MRI. Berrington de Gonzalez et al.[123] reported no net benefit of mammography surveillance before the age of 35 years in women with a *BRCA* mutation and recommended to limit the radiation dose by raising the age of mammography. Our results indicate that raising the age limit of supplemental mammography screening to the age of 40 years should be considered, not only for *BRCA* germline mutation carriers, but for all women at increased risk of developing breast cancer.

A further reason for this recommendation is that population-based mammography screening programs have been criticized because of overdiagnosis and overtreatment of non-fatal breast disease detected during screening[136]. Overdiagnosis, defined as the detection of a breast cancer at screening that would have never been identified clinically in the lifetime of the woman, has been reported between 1 - 10%[137]. Our results suggest that adding mammography screening to breast MRI may contribute to overdiagnosis because of the preferential detection of relatively indolent (pre-) malignant subtypes such as low grade calcified ductal in situ carcinoma as described in a previous study[125]. These cancers might be biologically irrelevant compared to invasive and in situ cancers detected with MRI that tend to be of higher grade and are usually detected at an earlier stage[75,135]. However, this is not evident from our data.

Our study has some limitations. It is a single center study in a tertiary referral center with a large high-risk screening program that might not be fully generalizable to the whole breast imaging community. In addition, due to the retrospective nature of the study, some of the exams of MRI and mammography were evaluated simultaneously which might affect the screening outcomes either positively or negatively. While the study describes a long time span the absolute number of cancers detected is still low, and more studies are required to confirm our findings.

In conclusion, mammography does not significantly add to cancer yield for all ages. Especially in younger women the added value is very limited and the number of mammography screens needed is very high. In higher age groups mammography does add to the detection rates, but also leads to an increase in FPR and FPB.
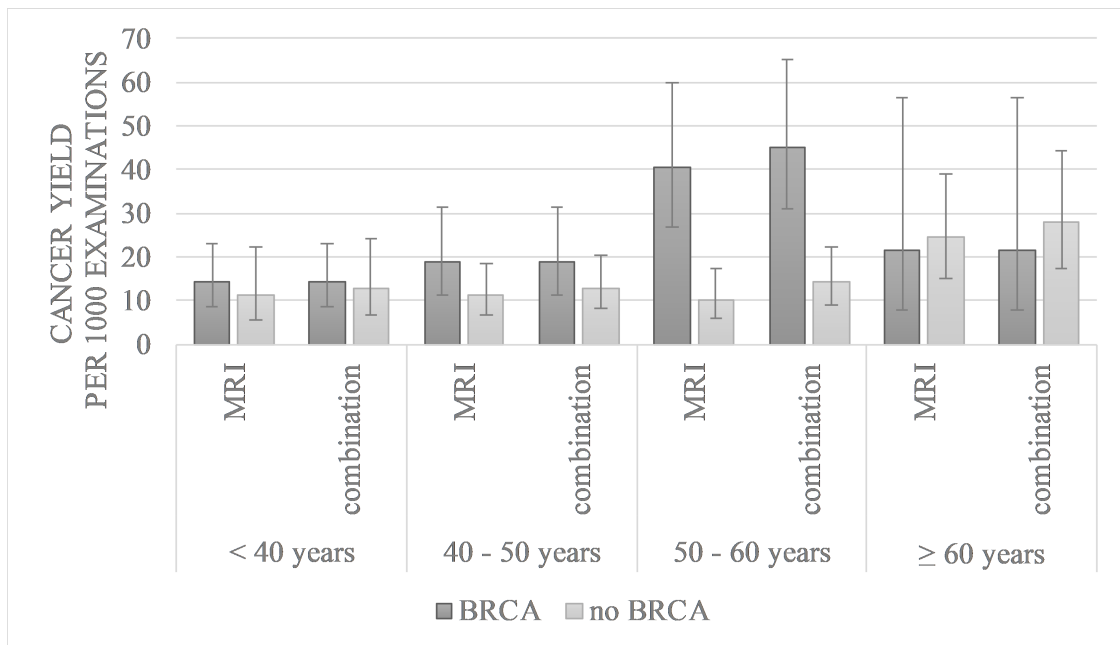

Figure 4.1: Cancer yield in women with a *BRCA* mutation versus all others (family, personal, and others). With the tag combination is meant the combination of mammography + MRI

Table 4.3: Number of screens needed (NSN) for one additional mammography-only detected cancer

	Age group (years)	Number of breast cancers	Number of screens	Breast cancers detected by mammography only	NSN for mammography to detect breast cancer missed by MRI
Overall	<40 years	26	1829	1	1829
	40-50 years	32	2050	2	1025
	50-60 years	45	1833	8	229
	≥60 years	22	841	2	421
BRCA	<40 years	17	1113	0	N/A
	40-50 years	14	737	0	N/A
	50-60 years	26	568	3	189
	≥60 years	4	190	0	N/A
No *BRCA*	<40 years	9	716	1	716
	40-50 years	18	1313	2	657
	50-60 years	19	1265	5	253
	≥60 years	18	651	2	326
Follow up	<40 years	17	1112	1	1112
	40-50 years	20	1447	2	724
	50-60 years	28	1342	8	168
	≥60 years	11	626	2	313
BRCA	<40 years	12	725	0	N/A
	40-50 years	9	554	0	N/A
	50-60 years	18	433	3	144
	≥60 years	0	152	0	N/A
No *BRCA*	<40 years	5	387	1	387
	40-50 years	11	893	2	447
	50-60 years	10	909	5	182
	≥60 years	11	474	2	237

N/A: not applicable, first round was not shown in the table since no mammography-only cancers were detected in first round

Table 4.4: Cancer yield, FPR and FPB results for mammography

Age category	Risk category	Mammography		
		Cancer yield* (95% CI)	FPR* (95% CI)	FPB* (95% CI)
<40 years	Overall	6.54 (3.73-11.46)	36.06 (27.57-47.03)	19.48 (13.73-27.58)
	BRCA	7.17 (3.61-14.18)	26.96 (18.07-40.03)	15.52 (9.03-26.57)
	no *BRCA*	5.61 (2.10-14.89)	49.62 (34.62-70.64)	24.94 (15.94-38.83)
40-50 years	Overall	6.35 (3.71-10.83)	40.44 (31.53-51.72)	24.23 (17.66-33.16)
	BRCA	6.78 (2.85-16.02)	18.90 (11.26-31.56)	12.21 (6.37-23.30)
	no *BRCA*	6.11 (3.08-12.06)	52.66 (39.80-69.38)	30.56 (21.36-43.55)
50-60 years	Overall	13.57 (9.21-19.94)	33.28 (24.94-44.29)	18.38 (12.60-26.75)
	BRCA	26.10 (15.93-42.47)	10.56 (4.75-23.31)	8.85 (3.67-21.15)
	no *BRCA*	7.89 (4.26-14.58)	43.06 (31.68-58.27)	22.16 (14.61-33.47)
≥60 years	Overall	17.72 (10.81-28.93)	25.82 (16.76-39.58)	10.88 (5.64-20.90)
	BRCA	21.48 (7.99-56.49)	10.69 (2.64-42.22)	0.00 (N/A)
	no *BRCA*	16.90 (9.54-29.77)	30.38 (19.27-47.57)	14.14 (7.31-27.15)
Overall	Overall	9.95 (7.80-12.69)	36.42 (31.42-42.19)	20.05 (16.49-24.37)
	BRCA	12.44 (8.75-17.65)	20.65 (15.32-27.77)	12.56 (8.50-18.51)
	no *BRCA*	8.33 (5.95-11.65)	47.08 (39.74-55.71)	24.95 (19.91-31.24)

* GEE used to calculate performance measures; correcting for multiple screening rounds within the same patient. All measurements are per 1000 exams; 95% CI: Wald 95% confidence intervals.

N/A: no cancers, recalls, or biopsies were found in this category and no 95% CI of this measure could be calculated.

Table 4.5: Cancer yield, FPR and FPB results for MRI

Age category	Risk category	MRI Cancer yield* (95% CI)	FPR* (95% CI)	FPB* (95% CI)
<40 years	Overall	13.09 (8.80-19.44)	87.25 (74.68-101.71)	62.30 (51.89-74.63)
	BRCA	14.31 (8.81-23.16)	75.49 (60.59-93.70)	56.72 (44.03-72.79)
	no *BRCA*	11.22 (5.61-22.31)	104.96 (84.38-129.84)	70.70 (54.44-91.35)
40-50 years	Overall	14.06 (9.83-20.07)	70.45 (59.44-83.31)	55.95 (46.39-67.33)
	BRCA	18.86 (11.26-31.42)	47.03 (34.06-64.61)	34.73 (23.93-50.15)
	no *BRCA*	11.37 (6.91-18.66)	82.49 (67.63-100.27)	67.55 (54.44-83.53)
50-60 years	Overall	19.56 (14.16-26.96)	41.64 (32.53-53.16)	29.88 (22.29-39.94)
	BRCA	40.37 (27.00-59.96)	44.36 (29.42-66.36)	30.66 (18.27-51.02)
	no *BRCA*	10.25 (5.97-17.55)	40.08 (29.42-54.38)	29.41 (20.59-41.84)
≥60 years	Overall	23.56 (15.39-35.93)	29.04 (19.45-43.15)	20.49 (12.72-32.85)
	BRCA	21.48 (7.99-56.49)	10.85 (2.66-43.12)	10.85 (2.66-43.12)
	no *BRCA*	24.40 (15.23-38.87)	34.37 (22.63-51.89)	23.31 (14.04-38.45)
Overall	Overall	16.64 (13.81-20.04)	62.52 (56.37-69.29)	46.27 (41.11-52.06)
	BRCA	22.08 (17.03-28.60)	56.98 (48.07-67.42)	41.81 (34.35-50.80)
	no *BRCA*	13.10 (10.02-17.10)	66.13 (58.03-75.27)	49.21 (42.42-57.03)

* GEE used to calculate performance measures; correcting for multiple screening rounds within the same patient. All measurements are per 1000 exams; 95% CI: Wald 95% confidence intervals.

N/A: no cancers, recalls, or biopsies were found in this category and no 95% CI of this measure could be calculated.

**Table 4.6:** Cancer yield, FPR and FPB results for the combination

Age category	Risk category	Combination Cancer yield* (95% CI)	FPR* (95% CI)	FPB* (95% CI)
<40 years	Overall	13.56 (9.24-20.10)	99.78 (86.35-115.04)	65.68 (54.96-78.31)
	BRCA	14.31 (8.81-23.16)	85.01 (69.45-103.66)	60.34 (47.29-76.70)
	no *BRCA*	12.65 (6.58-24.21)	121.64 (99.00-148.59)	73.85 (56.96-95.26)
40-50 years	Overall	15.02 (10.64-21.17)	89.74 (77.29-103.98)	63.26 (53.13-75.18)
	BRCA	18.86 (11.26-31.42)	57.42 (42.61-76.97)	42.39 (29.86-59.85)
	no *BRCA*	12.89 (8.10-20.48)	107.19 (90.27-126.84)	74.68 (61.11-90.97)
50-60 years	Overall	23.82 (17.82-31.76)	60.11 (48.79-73.85)	36.58 (28.05-47.59)
	BRCA	45.29 (31.14-65.43)	47.62 (32.20-69.90)	33.91 (20.91-54.54)
	no *BRCA*	14.16 (8.97-22.29)	65.53 (51.26-83.42)	37.51 (27.28-51.37)
≥60 years	Overall	26.49 (17.31-40.36)	46.86 (34.05-64.17)	24.16 (15.57-37.33)
	BRCA	21.48 (7.99-56.49)	22.05 (8.12-58.46)	10.85 (2.66-43.12)
	no *BRCA*	27.87 (17.39-44.38)	54.21 (38.71-75.43)	28.06 (17.67-44.30)
Overall	Overall	18.68 (15.65-22.27)	79.65 (72.70-87.21)	51.73 (46.28-57.78)
	BRCA	23.17 (18.00-29.79)	65.59 (56.09-76.55)	46.14 (38.31-55.47)
	no *BRCA*	15.74 (12.28-20.15)	89.14 (79.65-99.64)	55.39 (48.20-63.57)

* GEE used to calculate performance measures; correcting for multiple screening rounds within the same patient. All measurements are per 1000 exams; 95% CI: Wald 95% confidence intervals.

N/A: no cancers, recalls, or biopsies were found in this category and no 95% CI of this measure could be calculated.
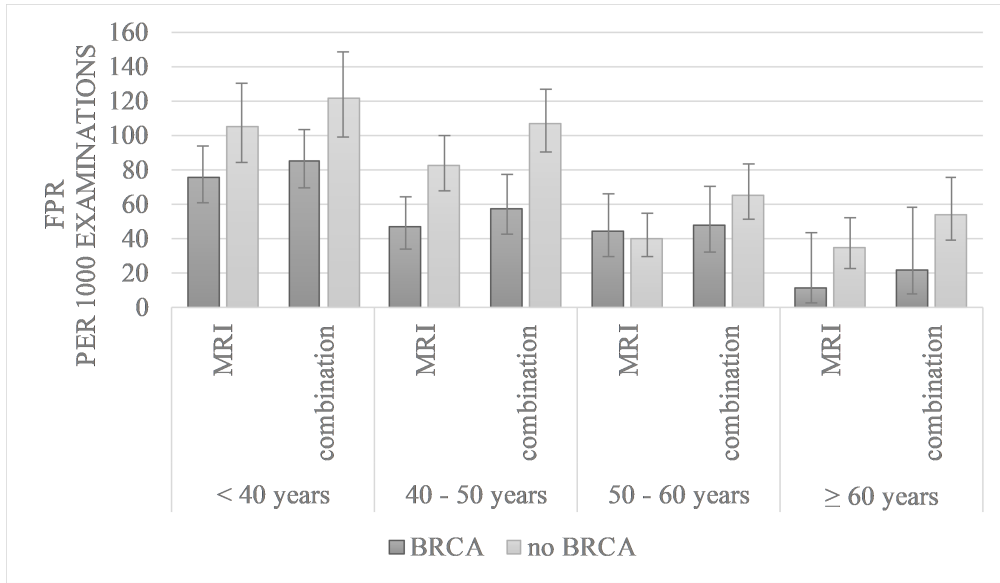
(a) False positives for recall
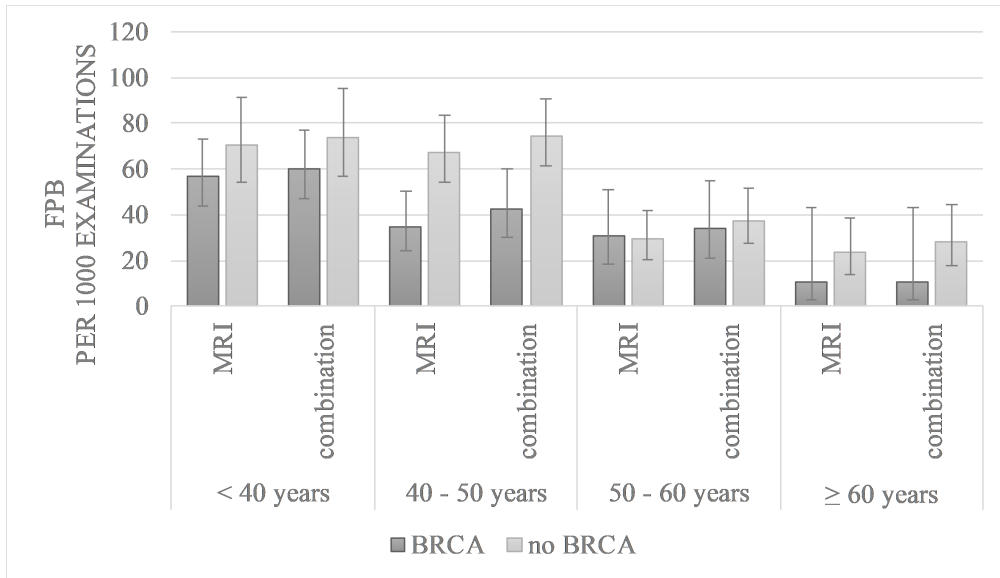
(b) False positives for biopsy

Figure 4.2: False positive rates for women with a *BRCA* mutation versus all others (family, personal, and others). With the tag combination is meant the combination of mammography + MRI

Visibility of breast cancers on prior MRI scans in breast MRI screening

5

Suzan Vreemann, Albert Gubern-Merida, Susanne Lardenoije, Peter Bult, Nico Karssemeijer, Katja Pinker, Ritse M. Mann

Abstract

Purpose To evaluate the frequency of missed cancers on breast MRI in women participating in a high-risk screening program.

Methods Patient-files from women who participated in an increased risk mammography and MRI screening program (2003 - 2014) were coupled to the Dutch National Cancer Registry. For each cancer detected, we determined whether an MRI scan was available (0 - 24 months before cancer detection), which was reported to be negative. These negative MRI scans were in consensus re-evaluated by two dedicated breast radiologists, with knowledge of the cancer location. Cancers were scored as invisible, minimal sign, or visible. Additionally, BI-RADS scores, background parenchymal enhancement, and image quality (IQ; perfect, sufficient, bad) were determined. Results were stratified by detection mode (mammography, MRI, interval cancers, or cancers in prophylactic mastectomies), and patient characteristics (presence of *BRCA* mutation, age, menopausal state).

Results Negative prior MRI scans were available for 131 breast cancers. Overall, 31% of cancers were visible at the initially negative MRI scan and 34% of cancers showed a minimal sign. Presence of a *BRCA* mutation strongly reduced the likelihood of visible findings in the last negative MRI (19 vs 46%, P < 0.001). Less than perfect IQ increased the likelihood of visible findings and minimal signs in the negative MRI (P = 0.021).

Conclusion This study shows that almost one-third of cancers detected in a high-risk screening program is already visible at the last negative MRI scan, and even more in women without *BRCA* mutations. Regular auditing and double reading for breast MRI screening are warranted.

5.1   Introduction

Breast Magnetic Resonance Imaging (MRI) is recognized as the most sensitive imaging method for the early detection of breast cancer[15]. Therefore, women at a lifetime risk for breast cancer development of $\geq$ 20% are invited for intensified screening programs including both mammography and breast MRI. Recent prospective studies have reported a sensitivity of MRI of around 90% using these screening settings[15,41,75,102,138].

Studies investigating the performance of mammography screening programs have consistently shown that between 31% and 50% of cancers detected in follow-up could have been detected at an earlier screening round[139–142]. This is one reason for the implementation of double reading in mammography screening[143,144]. Furthermore, the frequency of these errors in mammography screening is nowadays regularly audited.

Although recently auditing guidelines for breast MRI screening were published by the American College of Radiology[19], these do not include the evaluation of prior screening rounds in women who present with breast cancer. A few studies that focused on reasons and features of missed cancers in MRI screening showed that between 47% and 56% of cancers could have been detected at an earlier MRI examination[145,146]. However, to our knowledge, no studies report on the visibility of interval cancers, tumors detected on mammography or tumors in prophylactic mastectomy specimen in prior MRI scans.

Therefore, the purpose is to evaluate the frequency of visible breast cancers detected in women participating in a high-risk screening program, that were missed on a prior MRI scan (screen-detected, interval, and cancers in prophylactic mastectomy specimen), and to assess patient and imaging factors contributing to non-detection.

5.2   Materials and Methods

Ethics
The local institutional review board approved this study and the requirement for informed consent was waived.

Breast cancer screening in women at intermediate to high risk for breast cancer
The breast cancer screening program for women at high or intermediate risk ($\geq$ 20 -

25% lifetime risk) at our institution consists of annual breast MRI and mammography, usually at the same day, although some women opt for having the examinations 6 months apart. In women with germ-line *BRCA* mutations screening starts from the age of 25 years, and the first 5 years no mammography is performed. In other women at increased risk for breast cancer, screening starts from the age of 35 or 45.

All women underwent contrast-enhanced breast MRI in prone position using a dedicated breast coil on a 1.5T or 3T MRI scanner (Siemens, Erlangen, Germany). Contrast (Gd-DOTA, Guerbet, France) was administered through an iv-canula in the cubital vein at a dose of 0.1 mmol/kg using a power injector (Warrendale, Medrad, PA). Acquisition parameters changed over time, and are reported in detail elsewhere[105]. All MRI scans contained a 3D T1-weighted sequence that was performed before and 4 to 5 times after contrast administration. Clinical reporting was done by one of 8 board certified breast radiologists with between 6 months and 23 years of experience in breast MRI using a dedicated breast MRI workstation (versions of DynaCAD, Philips, Best, the Netherlands), that automatically yields subtraction images, maximum intensity projections (MIPs), and enhancement curves. All mammograms were obtained on full-field digital mammography machines (Senograph DS, Senograph 2000, GE, USA) by dedicated mammography technicians. Mammography was always performed in 2 planes (medio-lateral oblique and cranio-caudal), with optional further acquisitions. Clinical reporting using dedicated mammography monitors was done by the radiologists who also reported the MRI scan.

Case selection

All screening breast MRI and mammography exams performed from 01/01/2003 to 01/01/2014 were identified by a cross-computer search. This yielded 9571 screening MRI studies and 6553 mammograms obtained in 2773 women. This database was linked to the nationwide population-based Netherlands Cancer Registry. In our cohort, 164 women presented with a total of 179 cancers. For these women, we noted age, menopausal status, and the reason for intensified screening: *BRCA1*, *BRCA2*, family history of breast cancer, personal history of breast cancer, or other (including germline PTEN mutation, previous radiation to the chest, hormone replacement therapy, and lobular carcinoma in situ in an earlier biopsy).

Cancers (invasive cancers or ductal carcinoma in situ (DCIS)) were subsequently categorized into MRI-detected, mammography-detected, interval cancers, and cancers detected in prophylactic mastectomy specimen (further referred to as incidental cancers). For each of these categories, we obtained the images acquired at the time of

tumor detection and the last MRI reported to be negative. The time between cancer diagnosis and the last negative MRI scan was recorded. When no prior negative MRI was available, the case was excluded (N = 48).

MRI-detected cancers

Cancers were considered MRI-detected when they were screen-detected and mentioned in the MRI-report at the time of diagnosis (hereafter called the current MRI). The most recent breast MRI scan reported to be negative between 6 and 24 months before cancer diagnosis was selected for re-evaluation.

Mammography-detected cancers

Cancers were considered mammography-detected when they were screen-detected and described in the mammography-report, but the MRI-report was negative. In this case, the MRI performed within the same screening round as the mammogram was re-evaluated, and therefore the time between detection and the last negative MRI scan is negligible.

Interval cancers

Interval cancers were defined as cancers that were detected in between screening rounds due to symptoms. The last MRI prior to the cancer detection was selected.

Incidental cancers

Incidental cancers were defined as cancers detected in prophylactic mastectomy specimens with negative prior imaging. We selected the last MRI scan prior to the prophylactic mastectomy for assessment.

Retrospective MRI interpretation

The last negative breast MR-images were re-evaluated in consensus by two breast radiologists with respectively 8 and 12 years of experience in breast MRI. Readers were informed of the cancer location on the positive MRI and/or the histopathology results. In the case of mammography detected cancers, the location described in the mammography-report and histopathology results were given.

The review was performed on an in-house developed dedicated breast MRI workstation[147]. The workstation performed motion correction[148], and showed T1-weighted images, subtraction images and MIPs for all time-points. The average contrast enhancement versus time curve was shown for the pointer location. For the MRI-detected cases, the current MRI was displayed alongside the prior MRI.

The readers, in consensus, scored whether the cancer was either invisible, if there was a minimal sign, or the cancer was visible in the last negative MRI in analogy to the Dutch auditing practice for mammography screening[140]. When the MRI was truly negative, the cancer was rated as invisible. Minimal signs were visible lesions at the site of the later detected cancer that, according to the consensus reading, would not likely be recalled in screening practice. Visible lesions were lesions that were present at the site of the later detected cancer and should have been recalled according to the consensus reading. All lesions in the current MRI, as well as all lesions visible or showing a minimal sign in the prior MRI, were assessed according to the Breast Imaging Reporting and Data (BI-RADS) MR-lexicon and BI-RADS scores were given accordingly[19,149].

For each MRI scan, background parenchymal enhancement (BPE) was scored as minimal, mild, moderate, or marked. In addition, image quality (IQ) was subjectively scored as perfect, sufficient, or bad.

Performance measures
We first assessed the frequency of visible findings and minimal signs in the negative MRI scans, overall and in the subgroups (MRI-detected, mammography-detected, interval cancers, and incidental cancers). Subsequently, we investigated whether patient factors (age at cancer detection, menopausal state, the presence of a *BRCA* mutation) and imaging factors (field strength of MRI scanner, BPE, reported IQ) were related to the likelihood of missed lesions in prior breast MRI scans. For statistical analysis, we used Pearson's χ^2 tests for categorical variables and one-way ANOVA for continuous variables. The Tukey post-hoc test was used to compare the differences between the groups in case of continuous variables. To assess correlations, Spearman's rho was computed for ordinal values and when data were not normally distributed (time between scan and cancer detection). A two-sided P-value of ≤ 0.05 was considered statistically significant. All statistics were performed in SPSS (v22, SPSS Inc., Chicago, IL).

5.3   Results

Between 01/01/2003 and 01/01/2014, 131 breast cancers were detected in women participating in the intermediate and high-risk screening program, for whom a prior (or in case of mammography-detected cancers current) negative MRI scan was available (see Table 5.1). Of these cancers, 76 were MRI-detected, 13 mammography-

detected, 16 were interval cancers, and 26 were incidental cancers.

In the 131 re-evaluated MRI scans, lesions were considered invisible in 45 cases (34%). A minimal sign was present in 45 cases (34%) and lesions were visible in 41 cases (31%). Figure 8.1 presents examples of lesions in the three visibility cate-

Table 5.1: Time to negative MRI and time to cancer diagnosis (in months), stratified by detection mode and BI-RADS scores of prior MRI scans

	MRI-detect	MG-detect	Interval[1]	Incidental[2]	Overall
N (%)	76 (58)	13 (10)	16 (12)	26 (20)	131
Mean patient age in years (sd)	49.5 (11.2)	53.4 (8.3)	41.3 (9.1)	42.5 (10.6)	47.5 (11.2)
Time to last negative MRI in months (sd)	11.5 (3.0)	1.2 (3.4)	8.6 (3.1)	2.8 (3.1)	9.5 (4.5)
Histology					
IDC	49 (70)	2 (3)	15 (21)	4 (6)	70
ILC	4 (67)	0 (0)	1 (17)	1 (17)	6
DCIS	13 (33)	8 (21)	0 (0)	18 (46)	39
Mixed IDC & ILC	7 (78)	0 (0)	0 (0)	2 (22)	9
Other	3 (43)	3 (43)	0 (0)	1 (14)	7
Tumor stage at detection					
pTis	13 (33)	8 (21)	0 (0)	18 (46)	39
pT1mic	0 (0)	0 (0)	0 (0)	1 (100)	1
pT1a/b/c	7/22/18 (72)	3/1/0 (6)	0/1/7 (12)	2/3/1 (9)	12/27/26
pT2	12 (71)	1 (6)	3 (18)	1 (6)	17
pT3	1 (100)	0 (0)	0 (0)	0 (0)	1
pT4D	1 (100)	0 (0)	0 (0)	0 (0)	1
Recurrence	1 (33)	0 (0)	2 (67)	0 (0)	3
Unknown	1 (25)	0 (0)	3 (75)	0 (0)	4
Nodal status at detection					
pN0	52 (60)	11 (13)	7 (8)	17 (20)	87
pN+	16 (67)	1 (4)	5 (21)	1 (8)	24
Unknown	8 (40)	1 (1)	4 (20)	7 (35)	20
Visibility on prior					
Invisible	21 (47)	6 (13)	8 (18)	10 (22)	45
Minimal sign	31 (69)	4 (9)	3 (7)	7 (16)	45
Visible	24 (59)	3 (7)	5 (12)	9 (22)	41

[1] Interval cancers were defined as cancers detected in between screening rounds
[2] Incidental cancers were defined as cancers detected in prophylactic mastectomy specimens
  In between parenthesis the percentage of lesions in the specified category is given, except when indicated otherwise.

gories. Of all visible lesions, 2 (5%) were re-evaluated as BI-RADS 3, 35 (85%) as BI-RADS 4, and 4 (10%) as BI-RADS 5. Lesions that showed a minimal sign were scored as BI-RADS 2 in 22 cases (49%), and BI-RADS 3 in 23 cases (51%). Overall, 64 cases (49%) of prior negative MRI scans were scored as BI-RADS 3 or higher.

Table 5.2 shows the distribution of these findings in each subgroup. We did not observe a significant difference based upon the mode of cancer detection (P = 0.447). However, the frequency of visible findings in negative MRI scans of women with mammography-detected cancers was somewhat lower than in other categories, which is likely related to the simultaneous interpretation of mammograms and MRI scans in these cases and the fact that 8 of 13 cancers in this category were pure DCIS. It should be noted that 5 of 16 (31%) interval cancers were already visible at the last scan, and were all scored as BI-RADS 4 or 5.

Table 5.2 also shows the impact of patient and imaging factors on the likelihood of false negative MRI-reports. Overall, the frequency of missed findings is influenced by age, the reason for screening and reported image quality. Especially the presence of a *BRCA* mutation strongly reduces the frequency of visible findings in prior negative MRI scans.

5.4   Discussion

Our results show that, of 131 breast cancers with negative prior MRI exams retrospectively evaluated in this study, 31% were already visible on this negative exam. In fact, 29% of the 131 lesions were rated as BI-RADS 4 or BI-RADS 5 and should thus have been recalled based on these prior-exams. When including cancers presenting with minimal signs, 65% of the lesions were already recognizable on the prior MRI exam. Both from a learning perspective and in terms of liability it is essential that these figures are available.

These results are in line with reports from Yamaguchi et al.[145], and Pages et al.[146] who, in smaller cohorts (15 and 58 patients), reported that 56% and 47% of breast cancers were already visible on prior MRI examinations and retrospectively assessed as BI-RADS 3 or higher. Both studies, however, only included screen-detected cancers on MRI. To our knowledge, this study is the first to show that also 31% of interval cancers and even 35% of incidental cancers can be identified at the last negative MRI scan.

Table 5.2: Visibility of lesions on prior negative MRI scans, stratified by patient and imaging factors.

	N (%)	Invisible[1] (%)	Minimal sign[1] (%)	Visible[1] (%)	P-value[2]
Overall	131 (100)	45 (34)	45 (34)	41 (31)	N/A
Patient factors					
Mean age at cancer detection in years (sd)	47 (11.3)	44 (11.4)	46 (9.8)	52 (11.4)	0.003
Menopausal status					0.229
- Premenopausal	60 (46)	25 (42)	17 (28)	18 (30)	
- Postmenopausal	71 (54)	20 (28)	28 (39)	23 (32)	
Reason for screening					0.001
- *BRCA*[3]	70 (53)	34 (49)	23 (33)	13 (19)	
- non-*BRCA*	61 (47)	11 (18)	22 (36)	28 (46)	
Imaging factors					
Field strength					0.895
- 1.5T	96 (73)	33 (34)	32 (33)	31 (32)	
- 3T	35 (27)	12 (34)	13 (37)	10 (29)	
BPE					0.570
- Minimal	83 (63)	30 (36)	27 (33)	26 (31)	
- Mild	22 (17)	5 (23)	9 (41)	8 (36)	
- Moderate	12 (9)	3 (25)	4 (33)	5 (42)	
- Marked	14 (11)	7 (50)	5 (36)	2 (14)	
Reported IQ					0.021
- Perfect	111 (85)	43 (39)	34 (31)	34 (31)	
- Sufficient	18 (14)	1 (6)	10 (56)	7 (39)	
- Bad	2 (2)	1 (50)	1 (50)	0 (0)	

[1] According to the two readers in consensus in the prior MRI scan

[2] P-value is based on Pearson's χ^2 test evaluating the differences between the 3 visibility categories (invisible, minimal sign, and visible)

[3] The *BRCA* population included also untested first degree relatives

IQ=image quality

In between parenthesis the percentage of lesions in the specified category is given, except when indicated otherwise.

The program sensitivity of our high-risk screening cohort was, as reported earlier, 89.7%[125], which is comparable to recent prospective studies on MRI-screening in women at increased risk such as the Italian Hibcrit trial (91%)[41] and the German EVA trial (93%)[75]. Therefore, it is likely that our findings are applicable to all breast MRI screening settings. Our findings related to the visibility of cancers in prior-examinations are also similar to those found in mammography screening. Previous studies[140] have reported that up to 50% of the cancers detected in mammography

screening with double reading were already visible at earlier screening examinations and approximately half of these were suspicious.

The clinical consequences of missing a cancer in an MRI scan obtained before prophylactic mastectomy are relatively minor. Most of these cancers are DCIS only. This is a relatively frequent finding in prophylactic mastectomy specimen (approximately 5% in our institution)[110]. Since the period between the MRI scan and the subsequent prophylactic mastectomy is usually short, it is unlikely that missing these lesions
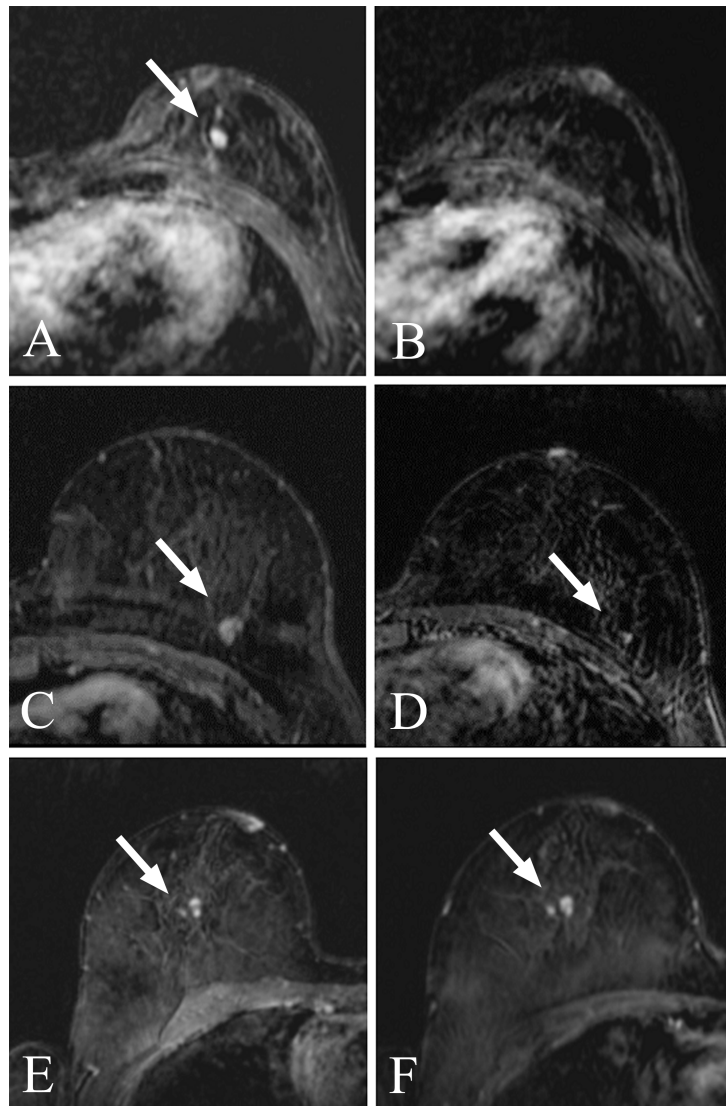
Figure 5.1: First row are the subtraction images of a breast cancer invasive ductal carcinoma (IDC) grade 3 (a), which were rated as invisible in the prior MRI (b), second row are the images (c) rated as minimal sign in the prior image (d) also showing an IDC grade 3, and the last row are the images of the current MRI (e) and the visible lesion in the prior MRI (f) an invasive lobular carcinoma grade 2.

will alter the patients prognosis. Nonetheless, the psychological impact might be substantial and should be taken into account[150].

The consequences of missing breast cancers that present subsequently as interval cancer, on the other hand, are dire. These cancers usually are invasive, poorly differentiated and fast-growing[151]. Detection of these cancers at the MRI-examinations could still have had a significant effect on subsequent prognosis and warrants investigation of methods to reduce false negative reporting of MRI scans.

In the group of mammography-detected cancers, the number of visible cancers is somewhat lower than in the other groups, though this did not reach statistical significance. This result was expected as the examinations are in practice often reported simultaneously. The fact that we did find 4 cancers on the current MRI scan that were re-evaluated as visible in retrospect can therefore only be explained by underreporting of subtle findings on the MRI scan due to already significant findings on the mammogram that would warrant a biopsy anyway. In addition, of note is that of 4 invasive cancers in this mammography category only one cancer was also regarded as invisible in the consensus reading. The finding that the frequency of invisible in situ cancers is substantial (50%) underlines that, although the sensitivity of breast MRI is high, it is not 100%, as we are certain these cancers are present in the breast.

The fact that the frequency of visibility of late(r) detected cancers is virtually independent of the eventual mode of cancer detection implies that the problem should be sought in the evaluation of MRI scans themselves. Radiologists, in general, may make two types of mistakes. The first is known as overlook error. The abnormality is simply not seen, and therefore cannot be classified correctly. The second is known as interpretation error. In this case, the lesion is seen but falsely interpreted as benign finding, and might therefore not even be reported, although the most evident cases of interpretation errors obviously mention the lesion as benign in the report[145,146,152,153]. In the evaluation of screening breast MRI, likely both processes play a role. In retrospect, 4 cases were classified as BI-RADS 5. It is unlikely that these lesions were seen but not recalled (see Figure 8.2). It is far more likely that these cases were overlooked. However, the vast majority (N = 35) of visible cases were classified in the prior-exam as BI-RADS 4. In these cases, interpretation errors might have prevented recall in clinical practice. This is also supported by the fact that the strongest modifier of the frequency of visible lesions is the presence of a *BRCA* mutation (49% of the lesions were truly invisible in *BRCA* mutation carriers compared to 18% in others). The much higher a priori chance of these women to
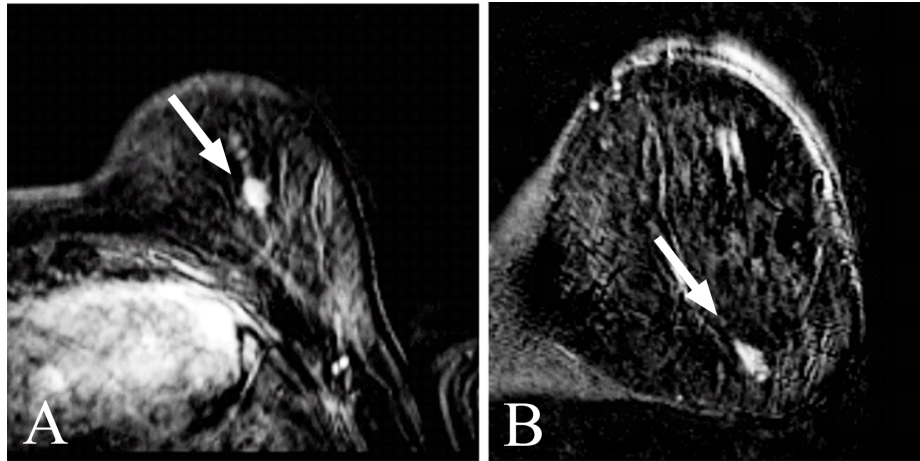
Figure 5.2: An example of an interval cancer detected on ultrasound nine months after prior MRI. This visible lesion was scored as BI-RADS 5 in re-evaluation; (a) subtraction image of an invasive ductal carcinoma grade 2 (pT1cN1mi (sn)) in transversal plane and (b) in coronal plane.

develop breast cancer compared to other women at increased risk[96], along with the fast growth of cancers in these women[154], leads to lowering the threshold for recall. In other words, women with *BRCA* mutations are recalled for lesions that would have been ignored in women without *BRCA* mutations. This might also explain the slightly younger age of women with invisible lesions at the last negative MRI, as the *BRCA* population, in general, is younger than the non-*BRCA* population.

We also observed that IQ was associated to the visibility score, with more frequent visible findings in negative reported MRI scans of less than perfect quality. This implies that striving for excellent image quality is important. When a scan is of inferior quality, for example, due to motion artefacts, rescanning the patient should be considered. We did not detect an effect of field strength on the visibility of lesions in prior negative MRI scans and, therefore, underline that adequate MRI scans can be performed at both 1.5 and 3T systems. BPE was not statistically related to the likelihood of visibility of lesions in the last negative MRI scan. Nevertheless, in women with marked BPE, the frequency of visible lesions was with 14% substantially lower than in other groups. We assume that very strong BPE might in fact sometimes obscure lesions that would otherwise be detected[155], though larger numbers are required for statistical analyses.

While an obvious solution to reduce misinterpretation of eventual cancers would be to lower the threshold for recall in non-*BRCA* mutation carriers, this cannot be straightforwardly applied. We earlier analyzed the positive predictive value of biopsy

(PPV3) in women in this cohort[125], which were 0.38 and 0.29 for *BRCA1* and *BRCA2* carriers, respectively. However, in women with a positive family history but no *BRCA* mutation PPV3 was only 0.14, which implies that lowering the threshold for recall in these women might lead to unacceptable high biopsy rates for benign lesions, and might further jeopardize the cost-effectiveness of MRI screening. The optimal balance between recall and cancer detection for women in different risk categories has yet to be determined.

We acknowledge that in the consensus read a bias was introduced by the non-blinded fashion of the evaluation. Readers were aware of the breast cancer location on the current MRI scan, mammography, or in the excised specimen. While this is common practice in auditing, it has been postulated that this approach is too harsh and leads to more lesions classified as visible than what would have been the case in a blinded setting where cases are mixed with normal scans[156]. However, in a previous study of Gubern-Merida et al.[132] a CAD system was used for the detection of breast cancer in a subset of 40 of the cancer cases that were used in the current evaluation, mixed with 120 normal cases. At 4 false positives per normal scan, the sensitivity of the CAD system was 0.71 and 0.31 for visible and minimally visible lesions in the prior negative scans, respectively, whereas sensitivity was 0.82 for the respective current scans in which the cancers were actually diagnosed. This implies that indeed a substantial subset of the cancers is at least detectable in the prior negative scans and patients could therefore indeed benefit from CAD-assisted reading or a blinded second read.

This study has some other limitations. It is a single center study based upon retrospective analysis of screening data. However, the fact that the program sensitivity is in line with published prospective trials implies that the findings are likely generalizable to similar bi-modal screening programs for women at increased risk. In addition, the prospective reads of the scans were conducted by one of eight radiologists with a strong variability (6 months to 23 years) in experience with breast MRI, which also reflects clinical practice. The numbers of missed cancers are too small to assess the effect of experience on false negative reporting. MRI protocols changed over time. However, all examinations were performed with at the time state-of-the-art equipment and imaging quality was in adherence with international recommendations[11,44]. Our numbers are likewise too small to detect whether MRI-protocol changes might have influenced the frequency of visible findings in negative MRI scans over time.

In conclusion, our results show that almost one-third of breast cancers can be detected at the last reported negative MRI. Since this is true for both screen-detected and interval cancers, it is essential to find methods to reduce reading errors. As a first step, regular auditing of clinical practice seems to be indicated. In addition, structural double reading of breast MRI-exams may be of value, although further research on the added cancer detection yield is required. Computer-aided detection tools for cancer detection in breast MRI might play a significant role in the prevention of reading errors in the future.

The association between BPE and tumor grade

Suzan Vreemann, Albert Gubern-Merida, Cristina Borelli, Peter Bult, Nico Karssemeijer, Ritse M. Mann

Abstract

Purpose Higher background parenchymal enhancement (BPE) could be used for stratification of MRI screening programs since it might be related to a higher breast cancer risk. Therefore, the purpose of this study is to correlate BPE to patient and tumor characteristics in women with unilateral MRI-screen detected breast cancer who participated in an intermediate and high-risk screening program. As BPE in the affected breast may be difficult to discern from enhancing cancer, we assumed that BPE in the contralateral breast is a representative measure for BPE in women with unilateral breast cancer.

Materials and Methods This retrospective study was approved by our local institutional board and a waiver for consent was granted. MR examinations of women with unilateral breast cancers screen-detected on breast MRI were evaluated by two readers. BPE in the contralateral breast was rated according to BI-RADS. Univariate analyses were performed to study associations. Observer variability was computed.

Results The analysis included 77 breast cancers in 76 patients (age: 48 ± 9.8 years), including 62 invasive and 15 pure ductal carcinoma in situ cases. A negative association between BPE and tumor grade ($P \leq 0.016$) and a positive association with progesterone status ($P \leq 0.021$) was found. The correlation was stronger when only considering invasive disease. Interreader agreement was substantial.

Conclusion Lower BPE in the contralateral breast in women with unilateral breast cancer might be associated with higher tumor grade and progesterone receptor negativity. Great care should be taken using BPE for stratification of patients to tailored screening programs.

6.1   Introduction

In dynamic contrast-enhanced breast magnetic resonance imaging (DCE-MRI), normal breast parenchyma may enhance after administration of a contrast agent. This enhancement is known as background parenchymal enhancement (BPE). The level of BPE after contrast administration is highly variable between women. Multiple factors, including age, pre- or postmenopausal status, phase in the menstrual cycle, and hormone usage can affect glandular tissue enhancement. Younger women have more often higher BPE and the degree of BPE naturally decreases with age[96].

Previous research showed that BPE may obscure or mimic lesion enhancement and can decrease the accuracy of breast MRI[157–161], even though contradictory results exist for its effect on sensitivity[159,162]. Nonetheless, lesion demarcation is impaired and it was shown that high BPE increases the likelihood of positive resection margins[163]. According to the newest BI-RADS MRI lexicon, BPE can be qualitatively evaluated rating the degree of enhancement as minimal, mild, moderate, or marked[19,164].

Recent literature focused more on the relationship between BPE and breast cancer occurrence. A study examining the relationship between breast cancer and BPE concluded that higher BPE is associated with a higher likelihood of breast cancer development[23]. Odds ratios increased from minimal to marked BPE by a factor of three to ten. This may have strong implications for personalized screening strategies, as these may be adapted to the level of BPE observed. This would be similar to using breast density in mammography as a parameter for stratification of women into more personalized screening programs. This is of particular interest in women at a lifetime risk of 20 to 50% because cost-effectiveness of MRI screening in these women is often doubted[165,166].

This raises the question whether high BPE is associated with the occurrence of all breast cancers or just a specific subset of cancer types, as this may affect the usefulness of BPE as a tool for stratification of women to more tailored screening programs. Since tumor size, tumor grade, hormone receptor status, and nodal status are the most essential parameters for long-term outcome prediction[167], it is essential to understand the correlation between BPE and these factors. However, only a few studies investigate the relationship between BPE in breast cancer patients and prognostic factors[168,169], and none evaluate cancers detected only through screening.

Therefore, the purpose of this study is to correlate BPE to patient and tumor char-

acteristics in women with unilateral breast cancer detected by a screening MRI examination who participated in an intermediate and high-risk screening program. As BPE in the affected breast may be difficult to discern from enhancing cancer, we assumed that BPE in the contralateral breast is a representative measure for BPE in women with unilateral breast cancer.

6.2   Materials and Methods

Screening program

This retrospective study was approved by our local institutional review board (CMO Arnhem-Nijmegen) and the requirement for informed consent was waived. The breast cancer screening program for women at increased breast cancer risk ($\geq$ 20-25% lifetime risk) at our institution consists of annual breast MRI in women aged from 25 to 60 in *BRCA* mutation carriers. In women of 30 years or older MRI is combined with mammography. In women at high familial risk, screening starts at 35 or 45 years combining MRI and mammography[20]. The examinations are generally acquired on the same day, although some women prefer to undergo mammography and breast MRI sequentially at six-month intervals.

Case selection

The local database of all breast MR imaging records was searched to identify all screening MR examinations performed between January 2003 and January 2014. Imaging data were cross-referenced with pathology records to identify all malignant lesions in this population. The inclusion criterion was histopathologically proven screen-detected breast cancer (invasive cancer or pure ductal carcinoma in situ). Screen-detected cancers were defined as cancers diagnosed after diagnostic workup initiated by screening findings. Women with a personal history of breast cancer, women who received radiation to the chest at a young age, and women who received hormone replacement therapy were excluded. Pathology records were reviewed to determine tumor characteristics according to the Dutch Guidelines for breast cancer[20]. Tumor characteristics were: histological type, histological grade (according to the modified Elston and Ellis criteria[167]), hormone receptor status (using immunohistochemistry for the estrogen receptor (ER) and progesterone receptor (PR) status and fluorescence in situ hybridization (FISH) for the human epidermal growth factor receptor 2 (HER2) status[170,171]), molecular subtype, primary tumor size (pT-stage) and lymph node status (pN-stage). Molecular subtype was defined based upon receptor status and proliferation markers, as described previously in more detail[172]. One year of follow-up was available for all patients.

Image acquisition

Breast DCE-MRI acquisitions were performed on either a 1.5 or 3 Tesla Siemens scanner (Magnetom Avanto, Magnetom Sonata, Magnetom Simphony or Magnetom Trio). All women were scanned in prone position using a dedicated bilateral breast coil. A transverse or coronal three-dimensional T1-weighted gradient-echo (GRE) dynamic sequence was performed before contrast agent administration followed by 4 or 5 post-contrast sequences. Subtraction series were created for all post-contrast time points. Motion correction was applied[148]. Pixel spacing (from 0.664 mm to 1.5 mm), slice thickness (from 1 mm to 1.5 mm), matrix (256 x 128, 448 x 381 or 512 x 96 pixels), echo time (from 1.71 msec to 4.76 msec), repetition time (from 4.56 msec to 8.41 msec) and flip angle (from 10° to 25°) differed among acquisitions because of the long time span of this study and the use of various scanners and protocols. Gadolinium-based contrast agents were administered at doses of 0.1 mmol/kg or 0.2 mmol/kg using a power injector (Medrad, Warrendale, PA) at a flow rate of 2.5 ml/s, followed by a saline flush. Premenopausal women were scheduled between the sixth and twelfth day of their menstrual cycle.

MR interpretation

All MR examinations of women with histopathologically proven unilateral breast cancer were reviewed by two experienced readers (a fifth year resident with experience in breast imaging (C.B.) and an experienced radiologist with nine years experience in breast MR imaging (R.M.M.)). The readers were informed of the location of the cancer but they were blinded to all other information. The two readers independently evaluated both the level of BPE in the contralateral breast and the level of motion for the complete volume, since motion might result in subtraction artefacts that might be mistaken for BPE. BPE was visually assessed according to the BI-RADS MRI lexicon as minimal, mild, moderate, or marked on the first post-contrast subtraction series obtained at approximately 90 seconds after contrast administration[19]. Motion was rated likewise on the same volumes as minimal, mild, moderate or severe.

Statistical analyses

We performed univariate analysis (χ^2 tests for categorical variables, Students T-test and one-way ANOVA for continuous variables) to investigate whether BPE was associated with certain patient and tumor characteristics (age, menopausal state, invasive versus in-situ disease, cancer type, tumor grade, ER status, PR status, HER2 status, molecular subtype, size of the primary tumor (pT-stage) and nodal status

(pN-stage)). In addition, we investigated the relation between BPE and scored motion in the same manner. To assess inter-reader variability, linear weighted kappa statistics (κ) were calculated. The strength of the kappa agreement was defined as <0.000 = poor, 0.000 - 0.200 = slight, 0.201 - 0.400 = fair, 0.401 - 0.600 = moderate, 0.601 - 0.800 = substantial and 0.801 - 1.000 = almost perfect.

Because the difference between minimal and mild, and moderate and marked is in clinical practice difficult to make, we chose to dichotomize BPE values in a second step. BPE was dichotomized into low BPE (original scores: minimal and mild) and high BPE (original scores: moderate and marked) to obtain more stable results. Separate analyses were performed for all included cancers and for invasive cancers only separately, and for cancers detected in *BRCA* patients and non-*BRCA* patients. A two-sided P-value of ≤ 0.05 was considered statistically significant. All statistics were performed in SPSS (version 22, SPSS Inc., Chicago, IL).

6.3   Results

In the period from January 2003 to January 2014, 10122 screening MR scans were performed in 2798 women. The cohort consisted of *BRCA* mutation carriers, women with a (strong) family history of breast cancer, women with a personal history of breast cancer, and women with other reasons for inclusion (including women with a germline PTEN mutation, women who had previous radiation to the chest at young age, and women who were diagnosed with lobular carcinoma in-situ in a previous biopsy). In total, 92 breast cancers (in 91 women) were screen-detected. Fifteen cancers in twelve women were excluded since these women had a personal history of breast cancer (N = 12), radiation to the chest (N = 2) or received hormone replacement therapy (N = 1). The final analysis included 77 cancers in 76 patients (median age of 48 years, range: 24 - 76 years). One woman had two primary breast cancers in the same breast (an invasive lobular cancer and an invasive ductal cancer).

Cancers

Patient and tumor characteristics are given in Table 6.1. The significant results of univariate analysis of patient and tumor characteristics in relation to BPE are presented in Table 6.2. BPE seems to be associated with tumor grade, scored motion, and to the PR status of the cancer. A substantial agreement in the assessment of BPE was found between R1 and R2 for all the cancers (κ = 0.719 (95% CI: 0.615 - 0.824)) and for invasive cancers only (κ = 0.750 (95% CI: 0.640 - 0.861)) using the original 4 categories. When using 2 categories the agreement was still substantial (κ = 0.633

(95% CI: 0.449 - 0.819)) for all cancers and for invasive cancers only (κ = 0.677 (95% CI: 0.488 - 0.865)). Figures 6.1 to 6.4 show examples of the four BPE categories.

Associations

For both readers there was a significant negative association between BPE and pathological tumor grade (R1; P = 0.016 and R2; P = 0.003 for all cancers, R1; P = 0.031 and R2; P = 0.007 for invasive cancers only, using dichotomized BPE scores, see Figure 6.5).

We investigated the association in *BRCA* mutation carriers and non-*BRCA* patients separately. The results show that there was no association between BPE and tumor grade in *BRCA* mutation carriers (P = 0.175 for both readers), however, there was a significant negative association in non-*BRCA* patients for reader 2 (P = 0.001).

We did also observe a significant positive association between BPE and motion in one reader (R1; P = 0.009), and a significant association between BPE and PR status for both the readers (R1; P = 0.018 and R2; P = 0.021, see Figure 6.6).
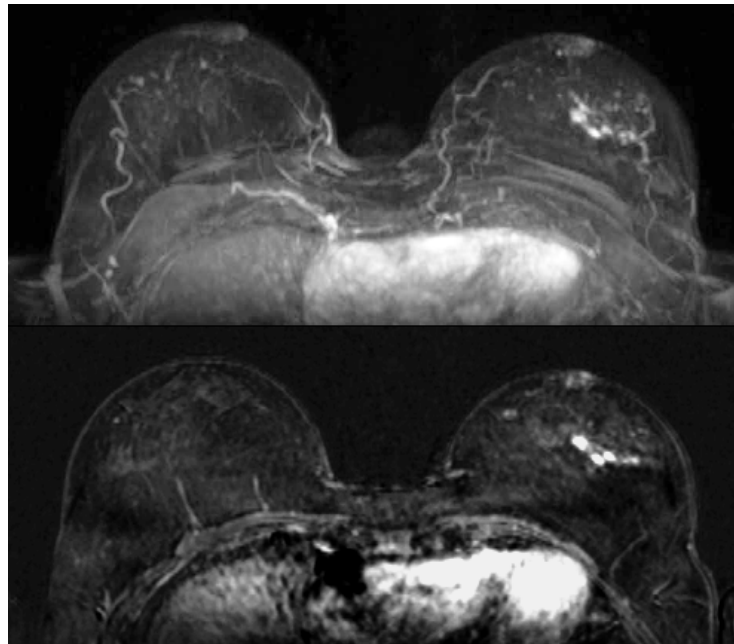

Figure 6.1: Maximum Intensity Projection (MIP, top image) and subtraction image (bottom image) of woman with unilateral cancer (59 years old, breast tumor in the left breast, IDC grade 2) with BPE rated as minimal by both readers.
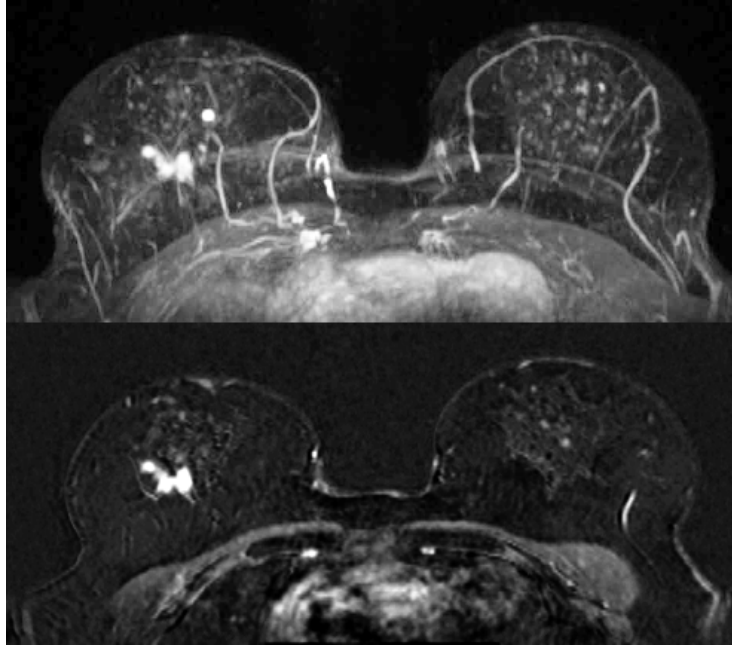
Figure 6.2: Maximum Intensity Projection (MIP, top image) and subtraction image (bottom image) of woman with unilateral cancer (61 years old, breast tumor in the right breast, DCIS grade 2) with BPE rated as mild by both readers.
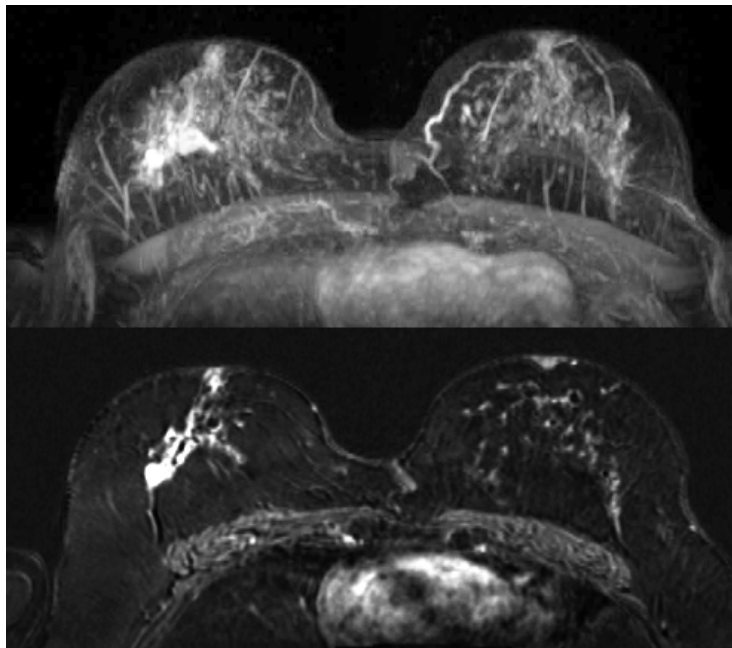

Figure 6.3: Maximum Intensity Projection (MIP, top image) and subtraction image (bottom image) of woman with unilateral cancer (70 years old, breast tumor in the right breast, IDC grade 2) with BPE rated as moderate by both readers.
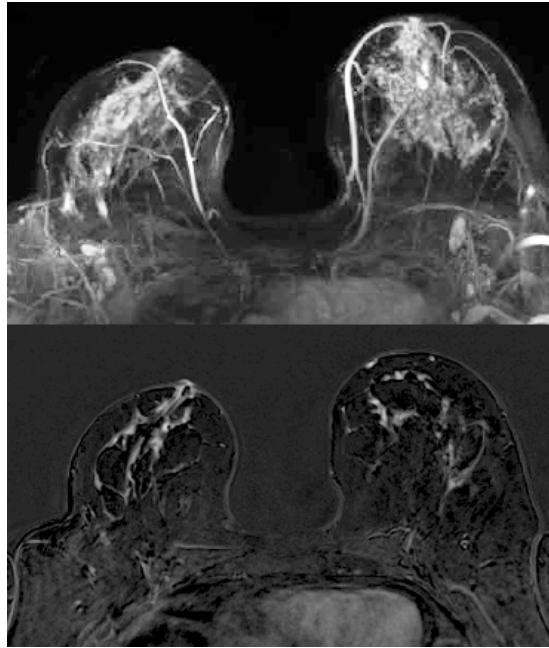
Figure 6.4: Maximum Intensity Projection (MIP, top image) and subtraction image (bottom image) of woman with unilateral cancer (60 years old, breast tumor in the left breast, IDC grade 1) with BPE rated as marked by both readers.
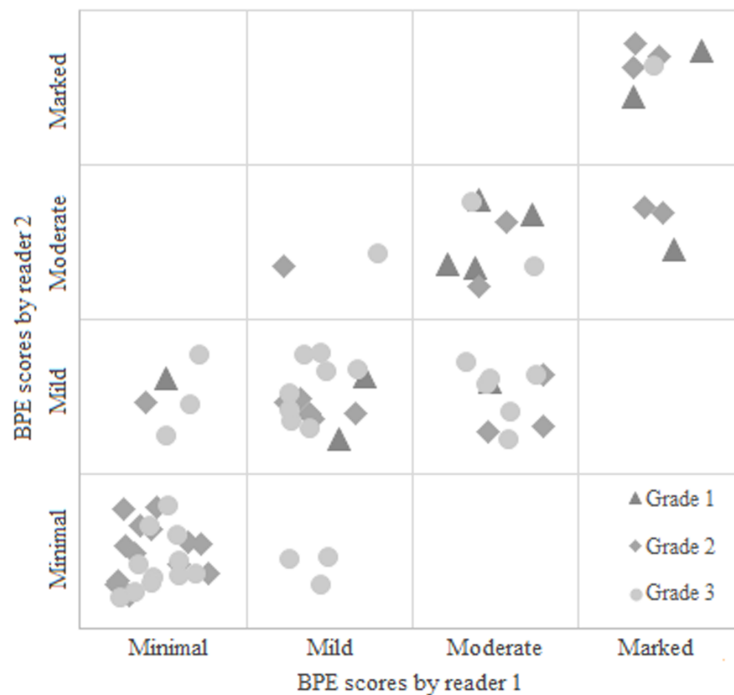

Figure 6.5: The association between BPE and tumor grade for Reader 1 and Reader 2.

6.4 Discussion

In this study, we investigated the association between BPE and patient and tumor characteristics in women at increased risk with screen-detected unilateral breast can-
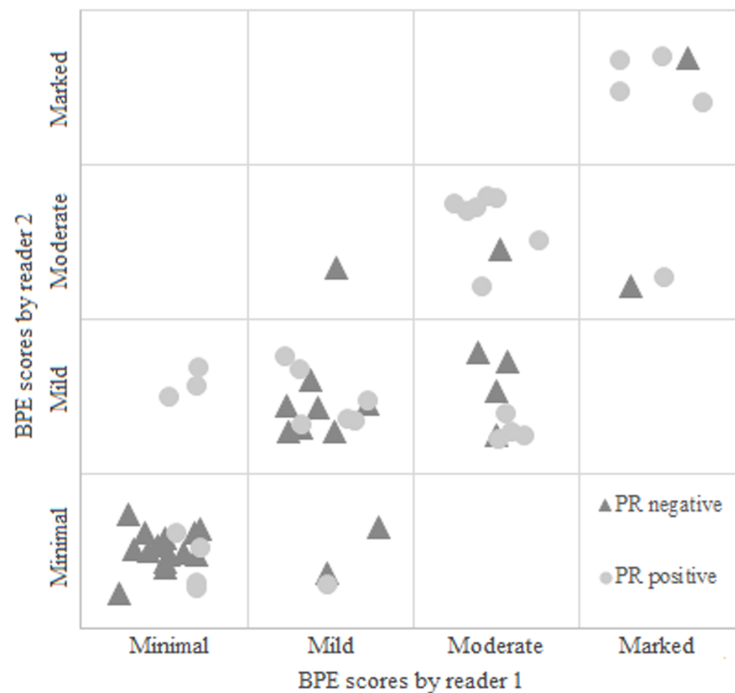
Figure 6.6: The association between BPE and PR status for Reader 1 and Reader 2

cer. Our results show that there is a negative association between BPE and tumor grade. Furthermore, a positive association between BPE and PR status is observed.

There is a worldwide intention to shift current research of breast cancer screening from population-based screening towards personalized screening. Based upon the results of King et al.[23] and the study of Telegrafo et al.[173], who detected a strong positive correlation between BPE and breast cancer risk, BPE has the potential to be used as a stratification factor for personalized breast screening, although this was not evident in all studies[159,174]. Considering the negative association between BPE and grade in our study, this would, however, mean that women who tend to develop low grade tumors might be more intensively screened than women who tend to develop high grade tumors. Similarly, the possible association between BPE and PR status should be taken into account.

The association between BPE and grade might be biologically explained by the fact that low grade cancers are in general also hormone receptor positive, whereas more aggressive cancers may be hormone receptor negative[175]. This might also explain the positive association observed between tumor PR status and BPE. It has been shown that increased estrogen and progesterone levels are related to an increased BPE[176,177], but also to increased breast cancer risk[178–183], especially for hormone receptor-positive

cancers[184], consequently relating high BPE to hormone receptor-positive cancers. However, we did not observe a correlation between ER status and BPE for any of the readers.

Our study also hints at a different pathogenesis for high grade tumors, in which hormonal stimulation is less important. This supports the theory that breast cancer development for high grade tumors is vastly different from low grade tumors[185,186]. This can also explain the difference between *BRCA* and non-*BRCA* patients (frequently sporadic tumors), as hormonal stimulation pathways might be different in *BRCA* tumor development, as described for the *BRCA1* gene by Hu et al.[187]. Our results may also partly explain the findings of van der Velden et al.[168]. The authors showed that parenchymal enhancement in the contralateral breast of women with invasive breast cancer is correlated with long-term outcome. Lower values of parenchymal enhancement showed potential as a predictive biomarker for relatively poor outcome in women who received endocrine therapy. This might be due to differences in tumor grade, but it could also be explained by the lower importance of the hormonal stimulation pathway for breast cancer growth in women with relatively low contralateral BPE. However, it must be taken into account that the definition of parenchymal enhancement of van der Velden et al. is different from the definition described in the BI-RADS MRI lexicon. In their study, automatically calculated late enhancement of the parenchyma (percentage of parenchymal enhancement over the whole post-contrast period) was used.

We found an association between BPE and increased motion scores in one reader (R1), which may be explained by the fact that motion can be misleading and might be incorrectly considered as BPE in subtraction images and vice versa[188,189]. However, further research is needed to investigate whether this holds true in a larger dataset and with more experienced readers.

Our findings are in contrast to the, to our knowledge, only study that so far investigated the relation between BPE and tumor characteristics[169]. Kim et al. reported that BPE of the contralateral breast is independent of tumor characteristics. Likely, this difference can be explained by the differences in patient cohorts. We only included women participating in an intermediate and high-risk screening program, whereas Kim et al.[169] included all patients with invasive ductal cancers who underwent preoperative MRI. The patients in our study had a different risk profile, although age (age under 50 years: 53.8% in current study versus 51.7% in the study of Kim et al.) and menopausal status (pre-menopausal: 45.0% in current study versus 47.9% in the

study of Kim et al.) between groups were comparable. We chose to study this population because these women are regularly screened using MRI and might benefit from personalized screening programs.

We calculated χ^2 tests to study associations between BPE and tumor characteristics. This was chosen over performing an analysis calculating odds ratios (including more variables as age and menopausal state) because tumor grade was the only prognostic factor that showed a significant relation to BPE in both of the readers, and the limited number of cases.

Our study has some limitations. Despite the fact that the number of cancers reported on in a high-risk screening program is relatively large, the absolute number of cancers included to study the correlations is relatively low. This holds particularly true for the number of pure DCIS cases (N = 15), although 19% pure DCIS (15/77) is a realistic representation of screening practice[190]. Furthermore, we suspect that the relatively low number of cancers might be one of the main reasons for not finding a significant association between BPE and age. Other studies showed a strong negative correlation of age with BPE[159,176]. This small sample size is also the main reason why no multivariable statistical methods are used in the current study. Future research is needed to test for possible confounders and interaction terms. In addition, continuous values of BPE could be used in future research to be able to define clear cut-offs.

Lastly, because of the long time span of this study, there is a large variability of MRI acquisitions in this dataset. Different MRI field strengths (1.5T and 3T), breast coils (4-channel, 7-channel, and 16-channel) and MRI protocols were used, and different types and amounts of contrast agents were injected. This may have strongly affected the amount of enhancement seen in the scans. However, the rating of BPE according to the BI-RADS MRI lexicon only considers the visual assessment of the fraction of fibroglandular tissue that enhances at 90 seconds after contrast administration and does not change based on differences in peak enhancement or wash-out pattern, which are more likely to be affected by the variability in scanning parameters. To our knowledge, only the study of Uematsu et al.[191] directly compared BPE in breast cancer patients at 1.5T and 3T and found no differences in the assessment between field strengths. The large variability, on the other hand, has the advantage that it reflects many of the breast MRI protocols currently in use and our findings, therefore, seem extendable to breast MRI screening in general.

In conclusion, BPE in the contralateral breast of patients with unilateral breast cancer in an intermediate and high-risk population might be negatively related to tumor grade and positively related to progesterone receptor status. Based on this finding, great care must be taken before using BPE as a method to stratify women at increased risk to more personalized MRI screening strategies. These results should, however, be validated in a larger study.

Table 6.1: Patient and cancer characteristics

Age	48 years +/- 9.9 years (range: 24 - 76 years)
Menopausal state	premenopausal: 36 (47%) postmenopausal: 41: (53%)
Cancer types	
DCIS	15 (19%)
Grade 1	0 (0%)
Grade 2	9 (60%)
Grade 3	6 (40%)
Invasive	62 (81%)
IDC	51 (82%)
ILC	10 (16%)
Other	1 (2%)
Grade 1	11 (18%)
Grade 2	22 (36%)
Grade 3	29 (47%)
ER status	Positive: 41 (66%)
	Negative: 20 (32%)
	Unknown: 1 (2%)
PR status	Positive: 30 (48%)
	Negative: 31 (50%)
	Unknown: 1 (2%)
HER2 status	Positive: 10 (16%)
	Negative: 50 (81%)
	Unknown: 2 (3%)
Molecular subtypes	Luminal A: 37 (60%)
	Luminal B: 3 (5%)
	HER2 type: 7 (11%)
	Basal-like: 13 (21%)
	Unknown: 2 (3%)
T-stage (invasive)	
Stage 1	42 (68%)
Stage 2	17 (27%)
Stage 3	1 (2%)
Unknown	2 (3%)
N-stage (invasive)	
Stage 0	38 (61%)
Stage 1	20 (32%)
Stage 2	1 (2%)
Stage 3	1 (2%)
Unknown	2 (3%)

DCIS: Ductal carcinoma in-situ, IDC: Invasive ductal carcinoma, ILC: Invasive lobular carcinoma, ER: Estrogen receptor, PR: Progesterone receptor, HER2: Human epidermal growth factor receptor 2

Table 6.2: Patient and cancer characteristics

Investigated parameters in association with BPE	Original scores BPE by R1	Original scores BPE by R2	P-value R1$	P-value R1#	P-value R2$	P-value R2#
Grade	grade 1/2/3	grade 1/2/3	0.040*	0.016*	0.008*	0.003*
Minimal BPE	1/14/14	0/13/14				
Mild BPE	2/7/12	4/10/17				
Moderate BPE	5/5/8	5/5/3				
Marked BPE	3/5/1	2/3/1				
Missing: 0						
PR status	PR+/PR-	PR+/PR-	0.089	0.018*	0.015*	0.021*
Minimal BPE	7/14	5/16				
Mild BPE	7/10	13/11				
Moderate BPE	11/5	8/3				
Marked BPE	5/2	4/1				
Missing: 17						
Scored motion per reader	minimal/mild/moderate/ and severe	minimal/mild/moderate/ and severe	0.038*	0.009*	0.031*	0.066
Minimal BPE	17/8/3/1	20/7/0/0				
Mild BPE	17/3/1/0	19/6/6/0				
Moderate BPE	4/7/5/2	7/6/0/0				
Marked BPE	4/3/2/0	4/2/0/0				
Missing: 0						

PR: Progesterone receptor

* is indicating a p-value $\leq$ 0.05.

$ P-values are based on original data (including the 4 BPE categories) using χ^2 tests

P-values are based on dichotomized values using χ^2 tests. Insignificant factors included: age (P $\geq$ 0.284 for original BPE values and P $\geq$ 0.188 for dichotomized BPE values), menopause (P $\geq$ 0.119 and P $\geq$ 0.055), invasiveness (P $\geq$ 0.383 and P $\geq$ 0.234), cancer type (P $\geq$ 0.284 and P $\geq$ 0.055), T-stage (P $\geq$ 0.242 and P $\geq$ 0.230), N-stage (P $\geq$ 0.356 and P $\geq$ 0.175), molecular subtype (P $\geq$ 0.165 and P $\geq$ 0.061), ER-status (P $\geq$ 0.119 and P $\geq$ 0.055), and HER2-status (P $\geq$ 0.700 and P = 0.999).

The association between FGT and BPE and breast cancer risk

7

Suzan Vreemann, Mehmet U. Dalmis, Peter Bult, Nico Karssemeijer, Mireille J.M. Broeders, Albert Gubern-Merida, Ritse M. Mann

Abstract

Background The amount of fibroglandular tissue (FGT) is known to be associated with an increased breast cancer risk. In recent case-control studies background parenchymal enhancement (BPE) on breast MRI was also proposed as a feature correlating to the development of breast cancer. In this study the predictive value of FGT and BPE for breast cancer development and risk of false positive findings, as measured at baseline, is investigated in women at increased risk for breast cancer.

Methods Negative baseline MRI scans of 1533 women participating in a screening program for women at increased risk for breast cancer between 01/01/2003 and 01/01/2014 were selected. Automated tools based on deep-learning were used to obtain quantitative measures of FGT and BPE. Logistic regression using forward selection was used to assess relationships between FGT, BPE, cancer detection, false positive recall and false positive biopsy.

Results Sixty cancers were detected in follow-up. FGT and BPE were not associated with breast cancer risk (P = 0.540 and P = 0.112, respectively). High FGT and BPE did lead to more false positive recalls at baseline (OR: 1.259, P = 0.050 and OR: 1.475, P = 0.003) and to more frequent false positive biopsies at baseline (OR: 1.315, P = 0.049 and OR: 1.807, P = 0.002), but were not predictive for false positive findings in subsequent screening rounds.

Conclusion FGT and BPE, measured on baseline MRI, are not predictive for breast cancer development in women at increased risk. High FGT and BPE lead to more false positive findings at baseline.

Impact FGT and BPE cannot be implemented in risk prediction models.

7.1   Introduction

Women at increased risk of breast cancer ($\geq$ 20 - 25% lifetime risk) are eligible for intensified screening programs, including a yearly breast Magnetic Resonance Imaging (MRI) study. Depending on the underlying risk factors, MRIs may be performed on an annual basis from the age of 25 (in *BRCA* mutation carriers)[11,12]. Women with a hereditary germline mutation and women with a history of radiation therapy to the chest at a young age are eligible for these programs. For other women, risk prediction tools are used to determine whether women are at increased risk and thus eligible for MRI screening. The current risk prediction tools rely mainly on personal factors, such as family history, age, race, etc.[192,193]. However, recent studies show that additional independent risk factors, including imaging biomarkers, might increase the predictive power of risk prediction.

Mammographic breast density (BD), for example, correlates to breast cancer risk in the general female population and in *BRCA* mutation carriers[194,195]. Consequently, a number of studies recommend adding BD to the available risk prediction tools[196–199].

The increased use of breast MRI allows for evaluation of additional risk factors to improve current risk prediction tools. Recent publications indicate that the amount of fibroglandular tissue (FGT) and/ or background parenchymal enhancement (BPE) measured on breast MRI may be useful to predict breast cancer risk in women undergoing breast MRI[23,24,162].

While in breast MRI all normal FGT enhances after contrast injection, the strength and speed of enhancement is dependent on variations in hormone levels, as determined by menstrual cycle phase, menopausal status, tamoxifen therapy, and hormone replacement therapy[158,200,201]. Studies of King et al. and Dontchos et al.[23,24] showed that higher amounts of BPE in the contralateral breast increase the risk of a breast cancer diagnosis. Their results thus suggest that BPE might be used for the prediction of breast cancer risk. Unfortunately, both studies evaluated BPE at time of breast cancer detection, and are therefore unable to document its predictive value for future breast cancer occurrence.

A further problem is that visual rating of BPE on a four-point scale (minimal: < 25%, mild: 25 - 50%, moderate: 50 - 75%, and marked: 75 - 100%), as used in studies so far, suffers from high interrater variability[53]. This limits its value for risk prediction. Analogue to the systems currently in use to automatically estimate BD on

digital mammograms, automated tools to assess FGT and BPE may reduce interrater variability and possibly provide more robust parameters for risk stratification.

The purpose of the current study is, therefore, to study whether FGT and BPE, as computed on a cancer-free baseline MRI-scan using an automated tool, are predictors of future breast cancer in an intermediate and high-risk breast cancer screening program. Furthermore, we evaluate whether FGT and BPE predict false positive findings.

7.2 Materials and Methods

This retrospective cohort study was approved by our local institutional review board and the requirement for informed consent was waived.

Screening program

The breast cancer screening program for women with a lifetime risk of $\geq$ 20% - 25% at our institution consists of annual breast MRI and mammography[11,20]. In *BRCA* mutation carriers, the screening regimen starts with breast MRI only at the age of 25. Mammography is added from age 30. Others start screening with both modalities at age 35 or 40, depending on the age at which relatives developed breast cancer. The first MRI scan performed for screening is hereafter referred to as the baseline MRI.

Case selection

The local database was searched to identify all patients who underwent breast MRI screening between 01/01/2003 - 01/01/2014. The case selection process is presented in Figure 7.1. Women of any age were included when they underwent at least two breast MRI examinations for screening in this period. We recorded for each patient whether a *BRCA* mutation was present and whether and when a risk-reducing salphingo-oophorectomy (RRSO) was performed. Women in whom a cancer was detected at the baseline MRI or within 6 months thereafter, women with a prior history of breast cancer, and women in whom the automated assessment of BPE failed were excluded. We did not exclude women who had a false positive finding in the first round of screening.

Ground truth

Normal or benign screening examinations were confirmed by at least one year of clinical follow-up and regarded as true negative when no cancer was detected before the subsequent screening examination. When no biopsy was indicated at short-
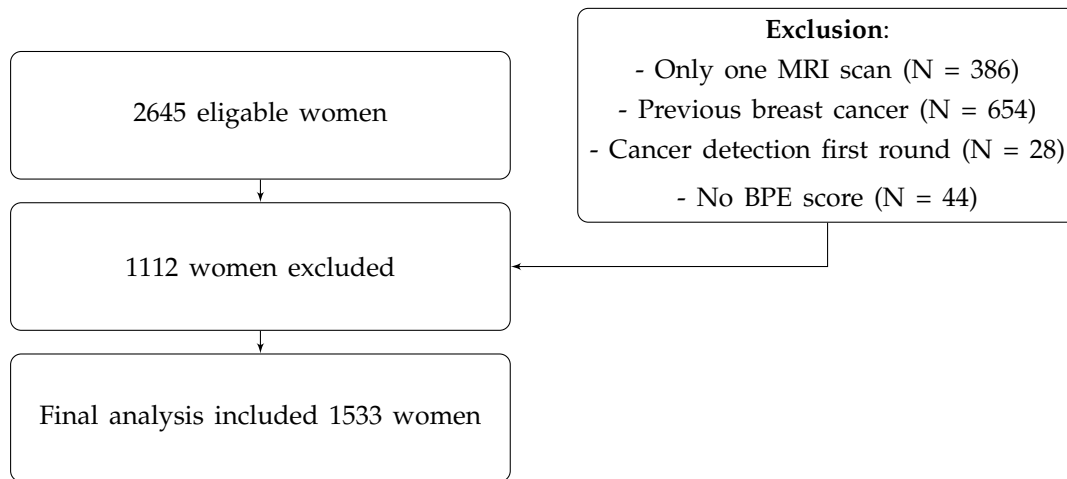
Figure 7.1: Flow diagram of the selection procedure.
BPE: background parenchymal enhancement

term follow-up, at least one year of clinical follow-up was required to confirm benignity. Biopsied lesions were identified by a cross-computer search with our pathology records. We subsequently analyzed if the biopsy was performed based on screening findings or for other reasons (e.g. symptoms).

Image acquisition

MRI Protocols varied over time[105]. Dynamic contrast-enhanced breast MRI acquisitions were performed on either a 1.5 or 3 Tesla Siemens scanner using a dedicated bilateral breast coil. Patients were placed in prone position. A transverse or coronal three-dimensional T1-weighted gradient-echo (GRE) dynamic sequence was performed before contrast agent administration followed by 4 or 5 post-contrast sequences. The first time point was acquired before intravenous agent injection and the following time points after injection of the contrast agent. The gadolinium chelates were administered at a dose of 0.1 mmol/kg or 0.2 mmol/kg using a power injector (Medrad, Warrendale, PA), followed by a saline flush.

Imaging interpretation

Automatic tools were used to objectively calculate percentages of FGT and BPE on breast MRI volumes. Breast and FGT were segmented on the native T1-weighted pre-contrast acquisitions using a deep-learning based method as described by Dalmiş et al.[105]. The fraction of FGT was calculated as the segmented volume of FGT divided by the total breast volume. BPE relative enhancement values were computed using the pre-contrast and the first post-contrast T1-weighted acquisition after motion correction[148], according to the ACR guidelines[19,126]. The fraction of BPE is expressed relative to the volume of the FGT, where an FGT voxel is considered to enhance if

it has a relative enhancement value higher than 10%. According to Dalmis et al.[202], who correlated BPE as scored by the automated tool to BPE as rated by radiologists, 10% relative enhancement correlates best. Figure 8.2 shows an example of automated computations of FGT and BPE. Final FGT and BPE measurements were the result of averaging over the two breasts of each woman. To verify whether correlations change when investigating different cut-off values, we performed the same analyses on 20%, 30%, 40% and 50% relative enhancement values (see Table 7.6).

Data-analysis

Women who developed cancer were identified by linkage of our data to the Netherlands Comprehensive Cancer Organisation. False positive MRI examinations were defined as examinations that led to recall in women in whom no breast cancer was detected. False positive recalls (FPR) include all women who were recalled (with or without the performance of biopsy). False positive biopsies (FPB) only include the women for whom the recall led to biopsy. We separated FPR and FPB in the first round from those that occurred in subsequent screening rounds.

Statistical analysis

Incomplete data was assumed to be missing at random and was excluded. Descriptive statistics were prepared with the use of contingency-table analyses for categorical data and Fisher's exact tests. The 95% confidence intervals (95% CI) for proportions were estimated using the Z-test for single proportions. Continuous data were compared with the Student's t-test or Pearson correlation coefficient (r) when normally distributed, otherwise, Mann-Whitney U tests were used. Bootstrapping (N = 1000) was used to calculate 95% CI. To increase statistical power, FGT and BPE were dichotomized into two categories based on the optimal categories in ROC-analysis (0 - 50th percentile and 50 - 100th percentile). A binary logistic regression model was constructed to find independent predictors for breast cancer or false positives. Separate and combined models were performed for FGT and BPE. The inclusion of variables in the model was based on existing knowledge of risk factors for breast cancer and/or false positives (covariates: age and *BRCA* status). Non-linear effects were evaluated using Box-Tidwell tests and when needed transformations were performed. The value of predictors was assessed by using forward feature selection (using a liberal probability-to-enter of 0.1). Interactions between predictors were evaluated in the final models by including interaction terms along with the main-effect terms. The final model was bootstrapped (N = 1000). Shrinkage using the heuristic method was applied to account for over-optimism[203]. Odds ratios (OR) were used to report on the relative odds of the occurrence of the outcome (future cancer, or false

positive result), where OR = 1 means that the predictor does not affect odds of outcome; OR > 1 means that the predictor is associated with higher odds of outcome; and OR < 1 means that the predictor is associated with lower odds of outcome. All statistical tests were two-sided. P ≤ 0.05 was considered significant. All statistics were performed in SPSS (version 22, SPSS Inc., Chicago, IL).

Table 7.1: Baseline patient characteristics

	Total cohort	*Developed cancer*	
	(N=1533)	*Yes (N=60)*	*No (N=1473)*
Age in years (median$, IQR)	41 (17.0)	40 (13.0)	42 (17.0)
BRCA mutation carriers (N, fraction)*	573, 0.37	41, 0.68	532, 0.36
FGT in percentage (median$, IQR)	12.7 (18.9)	11.6 (19.8)	12.7 (18.7)
BPE in percentage (median$, IQR)	67.7 (27.6)	71.3 (30.4)	67.6 (27.6)
Cancer (N, fraction)*	60, 0.04	60 (N/A)	N/A
- Age at cancer detection (median$, IQR)	42 (15.0)	42 (15.0)	N/A
False positive recall overall (N, fraction)*	337, 0.22	19, 0.32	318, 0.22
- Age at recall (median$, IQR)	42 (15.0)	40 (18.0)	42 (15.0)
False positive recall MRI (N, fraction)*	264, 0.17	16, 0.27	248, 0.17
- Age at recall (median$, IQR)	40 (15.0)	39 (16.8)	40 (15.0)
False positive biopsy overall (N, fraction)*	221, 0.14	12, 0.20	209, 0.14
- Age at biopsy (median$, IQR)	41 (14.5)	39 (16.5)	41 (14.5)
False positive biopsy MRI (N, fraction)*	203, 0.13	11, 0.18	192, 0.13
- Age at biopsy (median$, IQR)	40 (15.0)	38 (18.0)	40 (15.0)
RRSO (N, fraction)*	103, 0.07	5, 0.08	98, 0.07

N/A: not applicable, BPE: background parenchymal enhancement, FGT: amount of fibroglandular tissue, IQR: the difference between the 75th and 25th percentiles, RRSO: risk-reducing salpingo-oophorectomy

[*] fraction: fraction of positive cases

[$] Tested on normality using the Kolgomorov-Smirnov test

7.3 Results

Population

The final analysis evaluated baseline breast MRI scans of 1533 women, including 573 (37.4%) *BRCA* mutation carriers. Patient selection and exclusion is shown in Figure 7.1. The median age at baseline was 41 years (37 years for *BRCA* mutation carriers, and 44 years for others). In 60 (3.9%) women, cancer was identified after a negative baseline scan. 45 (75%) cancers were screen-detected cancers, 6 (10%) were interval cancers, and 9 (15%) cancers were detected in prophylactic mastectomies. 43 (71.7%) cancers were invasive and 17 (28.3%) were ductal carcinoma in situ (DCIS) only. The median time between the negative baseline scan and cancer detection was 3 years (2 in *BRCA* patients, 3 in others). Of the 573 *BRCA* mutation carriers 103 (18%) women

Table 7.2: Baseline characteristics of *BRCA* mutation carriers

| | *Total cohort* | *Developed cancer* | |
	(N=573)	*Yes (N=41)*	*No (N=532)*
Age in years (median$, IQR)	37 (17.0)	41 (14.5)	37 (18.0)
FGT in percentage (median$, IQR)	9.3 (14.5)	10.7 (16.7)	9.3 (14.5)
BPE in percentage (median$, IQR)	65.6 (26.7)	71.2 (33.3)	65.1 (26.1)
Cancer (N, fraction)*	41, 0.07	41, 1.00	0, 0.00
- Age at cancer detection (median$, IQR)	42 (14.5)	42 (14.5)	N/A
False positive recall overall (N, fraction)*	118, 0.21	12, 0.29	106, 0.20
- Age at recall (median$, IQR)	38.5 (15.0)	38 (19.5)	39 (14.3)
False positive recall MRI (N, fraction)*	97, 0.17	10, 0.24	87, 0.16
- Age at recall (median$, IQR)	38 (14.5)	38 (18.75)	39 (14.0)
False positive biopsy overall (N, fraction)*	80, 0.14	7, 0.17	73, 0.14
- Age at biopsy (median$, IQR)	38 (14.0)	34 (11.0)	39 (14.5)
False positive biopsy MRI (N, fraction)*	72, 0.13	7, 0.17	65, 0.12
- Age at biopsy (median$, IQR)	38 (15.3)	34 (11.0)	38 (15.5)
RRSO (N, fraction)*	103, 0.18	5, 0.12	98, 0.18

N/A: not applicable, BPE: background parenchymal enhancement, FGT: amount of fibroglandular tissue, IQR: the difference between the 75th and 25th percentiles, RRSO: risk-reducing salpingo-oophorectomy

* fraction: fraction of positive cases

$ Tested on normality using the Kolgomorov-Smirnov test

Table 7.3: Baseline characteristics of others at increased risk

| | *Total cohort* | *Developed cancer* | |
	(N=960)	*Yes (N=19)*	*No (N=941)*
Age in years (median$, IQR)	44 (15.0)	40 (11.0)	44 (15.0)
FGT in percentage (median$, IQR)	14.9 (20.7)	20.8 (20.5)	14.8 (20.6)
BPE in percentage (median$, IQR)	69.0 (27.6)	73.4 (29.1)	69.0 (27.6)
Cancer (N, fraction)*	19, 0.02	19, 1.00	N/A
- Age at cancer detection (median$, IQR)	43 (16.0)	43 (16.0)	N/A
False positive recall overall (N, fraction)*	219, 0.23	7, 0.37	212, 0.23
- Age at recall (median$, IQR)	43 (14.0)	46 (15.0)	43 (14.0)
False positive recall MRI (N, fraction)*	167, 0.23	6, 0.32	161, 0.17
- Age at recall (median$, IQR)	42 (15.0)	47.5 (17.8)	42 (14.0)
False positive biopsy overall (N, fraction)*	141, 0.15	5, 0.26	136, 0.14
- Age at biopsy (median$, IQR)	43 (14.0)	46 (10.0)	43 (14.0)
False positive biopsy MRI (N, fraction)*	131, 0.14	4, 0.21	127, 0.13
- Age at biopsy (median$, IQR)	42 (14.0)	47.5 (12.5)	42 (14.0)
RRSO (N, fraction)*	0, 0.00	0, 0.00	0, 0.00

N/A: not applicable, BPE: background parenchymal enhancement, FGT: amount of fibroglandular tissue, IQR: the difference between the 75th and 25th percentiles, RRSO: risk-reducing salpingo-oophorectomy

* fraction: fraction of positive cases

$ Tested on normality using the Kolgomorov-Smirnov test

had a RRSO prior to the first screening round.

337 (22.0%) women had a false positive recall. Seventy-three (21.7%) of these women were recalled based upon mammography findings. 264 (78.3%) women had at least

one false positive recall based on the MRI exam (total: 286 recalls on MRI), and 203 (13.2%) women had at least one false positive biopsy due to MRI findings (total: 217 biopsies). Median FGT measured on MRI was 12.7% (IQR: 18.9%), and median BPE was 67.7% (interquartile range (IQR): 27.6%). A more detailed presentation of the population characteristics is presented in Tables 7.1 - 7.3.

In univariate analysis, a significant association between FGT and *BRCA* status was found in both percentages (P = 0.001) and the dichotomous scores (P < 0.001). *BRCA* mutation carriers had lower FGT scores than others. Interestingly, the *BRCA* mutation carriers had a lower age at the baseline scan (median age of 37 for *BRCA* mutation carriers versus 44 for others, P < 0.001). A similar association was found between the percentage of BPE and *BRCA* status (P = 0.005), as *BRCA* mutation carriers had significantly lower BPE scores. When dichotomizing BPE, this remained significant (P = 0.020). FGT and BPE were negatively correlated to age (r = -0.289 and r = -0.129, P < 0.001), also when using dichotomous values (P ≤ 0.007). In *BRCA* mutation carriers coefficients were r = -0.418 (P < 0.001) and r = -0.132 (P = 0.002), respectively, and in women without a *BRCA* mutation r = -0.307 (P < 0.001) and r = -0.152 (P < 0.001). FGT and BPE were not correlated (P = 0.879). Plots of the univariate analysis are presented in Figures 7.3 - 7.10. In *BRCA* mutation carriers, BPE was not associated with a history of RRSO (P = 0.886, Table 7.2).

Cancer prediction models

In univariate analysis, FGT was not associated with breast cancer for both discrete (P = 0.768) and dichotomous values (P = 0.511). In regression analysis, FGT was not considered an independent risk factor for breast cancer, only *BRCA* status was (OR: 3.615, P = 0.001). Likewise, percentages and dichotomized BPE scores of the baseline MRI scan were not associated with breast cancer (P = 0.625 and P = 0.236, respectively). In regression analysis, adjusting for the only significant risk factor (*BRCA* status), BPE was also no significant predictor of cancer (P = 0.112). When evaluating both FGT and BPE, both were not significantly associated with breast cancer risk (P = 0.824 and P = 0.112). Also in the subgroup of the *BRCA* and non-*BRCA* mutation carriers only, BPE and FGT were not associated to breast cancer. Details of predictors can be found in Table 7.4.

False positive recall (FPR) models

When investigating the first round results alone (diagnostic model), both FGT and BPE were correlated to higher FPR rates (OR: 1.259; P = 0.050, and OR: 1.475; P = 0.003, respectively). For subsequent rounds (prognostic model), higher FGT at

baseline was still significantly related to higher FPR in both continuous and dichotomized values (P $\leq$ 0.029). BPE, however, was not related to FPR (P $\geq$ 0.818) in univariate analysis. In regression analysis only age remained as related factor to FPR in follow-up (OR: 0.955, P = 0.001, Table 7.5).

False positive biopsy (FPB) models

When only investigating the first round (diagnostic model), both FGT and BPE were correlated to higher FPB (OR: 1.315 (P = 0.049) and OR: 1.807 (P = 0.002), respectively).When excluding the FPB in the first round (prognostic model), FGT and BPE were both not related to false positive biopsies (P $\geq$ 0.066) in univariate analysis. Regression analysis showed that age was negatively related to FPB in follow-up (P = 0.001, Table 7.5).

No interaction terms were found to be significant in none of the prediction models. In addition, changing levels of BPE cut-offs did not change any of the conclusions for both the cancer- and false positive prediction models (Table 7.6 and 7.7).

Table 7.4: Regression coefficients and odds ratios for the prognostic cancer prediction model

Model	Predictor	P-value	Included in final model	β (95% CI)	OR (95% CI)	Shrinkage factor
Overall population						
Cancer-FGT	*BRCA**	<0.001	x	1.285 (0.762 - 1.872)	3.615 (2.143 - 6.501)	0.96
	Age	0.930	-			
	FGT	0.511	-			
Cancer-BPE	*BRCA**	<0.001	x	1.285 (0.769 - 1.875)	3.615 (2.158 - 6.521)	0.96
	Age	0.930	-			
	BPE	0.236	-			
Subgroup *BRCA*						
Cancer-FGT	Age	0.330	-	N/A		
	FGT	0.936	-	N/A		
Cancer-BPE	Age	0.330	-	N/A		
	BPE	0.106	-	N/A		
Subgroup non-*BRCA*						
Cancer-FGT	Age	0.126	-	N/A		
	FGT	0.621	-	N/A		
Cancer-BPE	Age	0.126	-	N/A		
	BPE	0.641	-	N/A		

For every model different shrinkage factors were used, shrunk β and OR are presented.

β: standardized coefficients, OR: odds ratio, CI: confidence interval, FGT: amount of fibroglandular tissue, BPE: background parenchymal enhancement, N/A: not applicable

* *BRCA* = 0 is reference category

7.4 Discussion

The purpose of this study was to investigate the predictive value of FGT and BPE in predicting breast cancer risk in a population at increased risk of developing breast cancer. Additionally, the effect of FGT and BPE on false positive recalls and biopsies was investigated. Our results show that neither FGT nor BPE at baseline was associated with the risk of developing breast cancer. Both higher FGT and BPE did lead to

Table 7.5: Regression coefficients and odds ratios for the effect on current and subsequent MRI scans on false positive findings

Model	Predictor	P-value	Included in final model	β (95% CI)	OR (95% CI)	Shrinkage factor
Diagnostic model for false positive findings (current MRI scans)						
FP Recall-FGT	*BRCA*	0.771	-			
	Age	0.167	-			
	FGT*	0.050	x	0.230 (0.012 - 0.451)	1.259 (1.012 - 1.569)	0.74
FP Recall-BPE	*BRCA*	0.612	-			
	Age	0.102	-			
	BPE*	0.003	x	0.389 (0.120 - 0.666)	1.475 (1.128 - 1.946)	0.87
FP Recall-FGT and BPE	*BRCA*	0.894	-			
	Age	0.229	-			
	FGT*	0.072	x	0.251 (-0.013 - 0.535)	1.285 (0.987 - 1.707)	0.83
	BPE*	0.005	x	0.366 (0.111 - 0.625)	1.442 (1.118 - 1.868)	
FP Biopsy-FGT	*BRCA*	0.350	-			
	Age	0.496	-			
	FGT*	0.049	x	0.274 (0.022 - 0.568)	1.315 (1.022 - 1.765)	0.74
FP Biopsy-BPE	*BRCA*	0.269	-			
	Age	0.362	-			
	BPE*	0.002	x	0.592 (0.253 - 0.957)	1.807 (1.288 - 2.605)	0.91
FP Biopsy-FGT and BPE	*BRCA*	0.453	-			
	Age	0.651	-			
	FGT*	0.064	x	0.312 (-0.012 - 0.677)	1.367 (0.988 - 1.968)	0.87
	BPE*	0.002	x	0.559 (0.218 - 0.911)	1.750 (1.243 - 2.487)	
Prognostic model for false positive findings						
FP Recall-FGT	*BRCA*	0.773	-			
	Age	0.001	x	-0.047 (-0.069 - -0.026)	0.955 (0.933 0.975)	0.95
	FGT*	0.224	-			
FP Recall-BPE	*BRCA*	0.773	-			
	Age	0.001	x	-0.047 (-0.069 - -0.026)	0.955 (0.933 0.975)	0.95
	BPE*	0.932	-			
FP Biopsy-FGT	*BRCA*	0.892	-			
	Age	0.001	x	-0.051 (-0.081 - -0.026)	0.951 (0.922 0.974)	0.94
	FGT*	0.557	-			
FP Biopsy-BPE	*BRCA*	0.892	-			
	Age	0.001	x	-0.051 (-0.081 - -0.026)	0.951 (0.922 0.974)	0.94
	BPE*	0.572	-			

For every model different shrinkage factors were used, shrunk β and OR are presented. β: standardized coefficients, OR: odds ratio, CI: confidence interval, FGT: amount of fibroglandular tissue, BPE: background parenchymal enhancement

* FGT and BPE = low is reference category

higher odds ratios for false positive findings in the baseline examination. We did not observe any predictive value of FGT or BPE for FPR or FPB in subsequent screening rounds.

It has already been well established that mammographic BD impairs mammographic sensitivity[204]. In an average risk population, BD is also known to correlate with breast cancer risk[194]. In line with the studies from Dontchos et al.[24] and Passaperuma et al.[205] who reported that neither mammographic BD nor FGT on MR images was predictive of breast cancer risk in women at increased risk, we did not observe a correlation between the fraction of FGT and the development of breast cancer in our high risk cohort. However, Mitchell et al.[195] reported contradictory results. In their study it was suggested that high BD in *BRCA* mutation carriers increased the risk of breast cancer, with a relative risk similar to that observed in the general population. The discordance between this report and our findings may at least partly be explained by the automated volumetric FGT estimation in our study, which provides a different representation of the FGT than visually inspected BD in mammography, albeit previous studies showed a clear correlation between these measurements[206]. It may also be related to the limited sample size in our study and other studies published thus far.

Current clinical practice is shifting towards personalized screening, making risk prediction tools increasingly important. Recent case-control studies have shown that BPE might be predictive of breast cancer risk[23,24], although contradictory results exist for non-high risk women[207]. However, in these studies the BPE scores of the healthy breast (partly for Dontchos et al.[24]) in breast cancer patients were compared to BPE scores in healthy controls. The current study, in which the BPE before cancer development is evaluated in actual patients, suggests that BPE is not predictive for breast cancer in women at increased risk. A possible explanation for this is that in case-control studies BPE in cancer patients might have been affected by the presence of breast cancer. Consequently, further research into the biological basis and modifying factors of BPE is needed. Alternatively, our results might point to a different carcinogenesis in women at increased risk.

Evidence suggests that BPE correlates negatively with age and increases with hormonal activity[176,208,209]. Interestingly, our results showed that *BRCA* mutation carriers had significantly lower FGT and BPE values compared to women without a *BRCA* mutation, while the age of *BRCA* mutation carriers was significantly lower than that of women without *BRCA* mutation. This counterintuitive result may be

due to differences in the effect of hormones on FGT in women with and without *BRCA* mutation[210]. The fact that we did not observe a difference in BPE between the *BRCA* mutation carriers who did and did-not undergo a RRSO before the baseline MRI (P = 0.886) also points in this direction. Nevertheless, prior research showed that RRSO may still reduce both BPE and FGT values[211], and therefore our results need to be interpreted with caution as they might also be explained by the relatively low number of women who underwent RRSO in our study.

Women with high BPE scores had a 1.5 times higher chance to get a FPR, and 1.8 times higher chance to get a FPB in the first screening round. This is in line with previous studies, describing that more focal, regional, or asymmetric BPE was associated with a higher likelihood of BI-RADS 3 assessment in the screening setting[174]. Giess et al. stated that, in the latter case, it may be hard to distinguish BPE from non-mass enhancement (NME)[212]. Consequently, when the enhancement pattern is interpreted as NME, the reporting radiologist has to consider the possibility of malignancy, thus chances on false positives increase. DeMartini et al.[159] also reported that higher amounts of BPE were associated with higher rates of abnormal interpretation. Brennan et al. reported that moderate and marked BPE are associated with significantly higher MR imaging-guided core biopsy cancellation rates compared to minimal or mild BPE[213]. However, the even stronger correlation between BPE and FPB in our study unfortunately suggests that many biopsies are still performed due to BPE. Nevertheless, neither BPE nor FGT is predictive of false positive recalls or biopsies in subsequent screening rounds, which could mean that BPE and FGT are only affecting false positives when no prior exams are available.

The automated algorithm for BPE estimation eliminates intra- and interrater variability. This is relevant as previous studies reported only a fair interrater variability for BPE when using observer scores according to the BI-RADS lexicon[53]. The automated method provides quantitative measurements and therefore creates an opportunity to define more precise cut-off points. The chosen cut-off was selected based upon previous research, but it is possible that different cut-off points might lead to different results, but did not lead to different conclusions. This is in line with a recent study on the prognostic value of BPE in the contralateral breast of women with unilateral breast cancer, where the effect of different cut-offs appeared to be minimal[214]. Nevertheless, it should be noted that the different methods to assess BPE may also lead to different outcomes in the risk model.

Our study has some limitations. Due to the retrospective nature of our data, it was

not always possible to retrieve data on the menstrual cycle or menopausal status. Therefore, we could not correct for these factors. In addition, this was a single institutional study, which potentially limits its generalizability. During the study period, we changed from a 1.5 T scanner to a 3 T scanner, and also adapted the scanning protocol several times which could potentially influence the results of the BPE calculation algorithm. Another possible limitation of the study was that, only in the case of false positive findings, we did not exclude women who had a false positive finding directly after the first screening round. In theory this could alter FGT and BPE scores, although we minimized this effect by averaging scores over two breasts.

In conclusion, automatically computed FGT and BPE measures at baseline were not associated with subsequent breast cancer occurrence in a cohort of women at high risk for breast cancer. This has implications for personalized screening, as FGT and BPE cannot be implemented in risk prediction models. Higher FGT and BPE were, however, associated with higher rates of false positive findings at baseline, patient counseling should therefore include these outcomes before starting MRI screening.
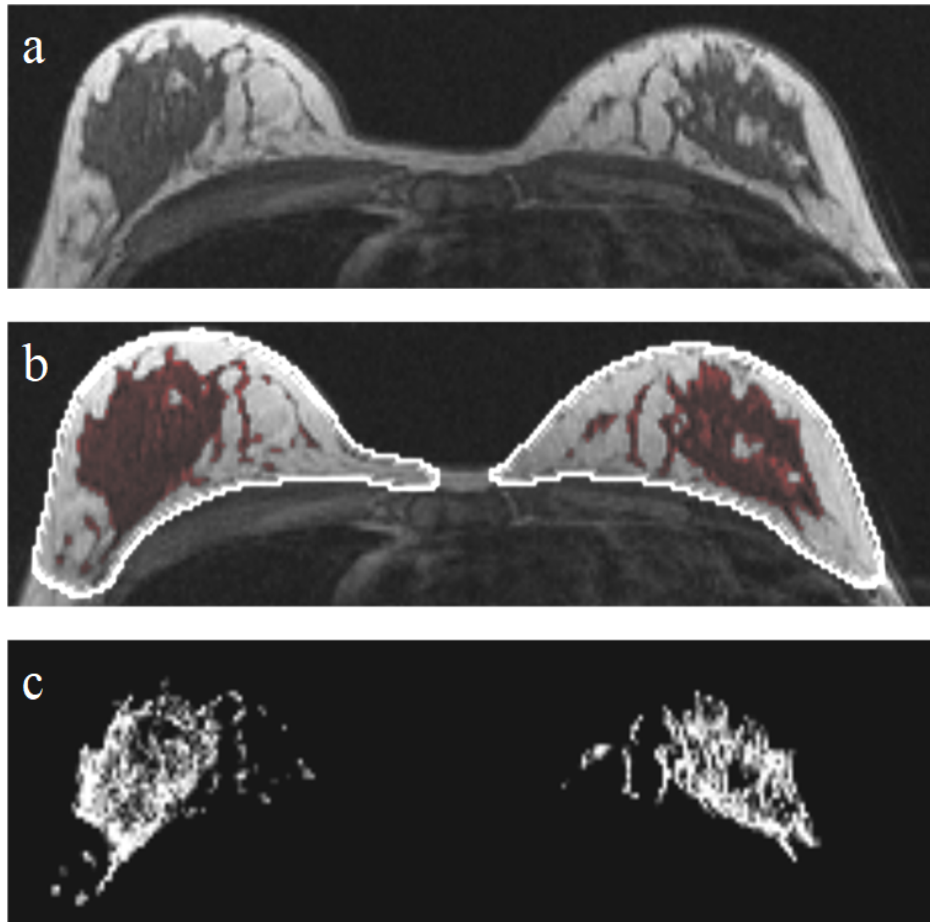
Figure 7.2: An example of the steps of the automated tool. First, the original image (a), then the breasts and parenchymal tissue are segmented (b), and finally, relative enhancement values are computed for the segmented FGT volumes (c). BPE values are extracted from the enhancing voxels within the parenchymal tissue, based on these relative enhancement values.

Table 7.6: BPE values for different cut-off values of the deep-learning system

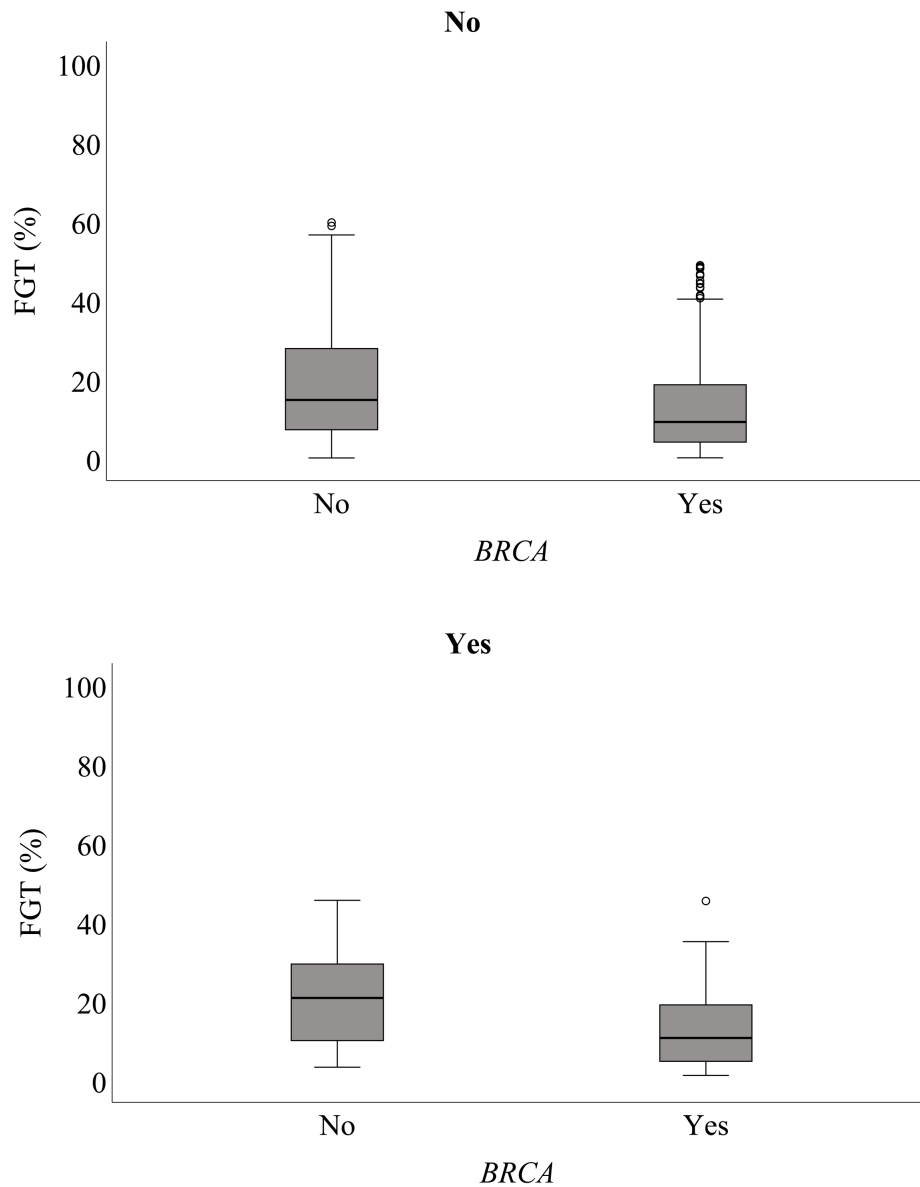|	*Total cohort*	*Developed cancer*	
		Yes	*No*
BPE in percentage (median, IQR) (10%)	67.7 (27.6)	71.3 (30.4)	67.6 (27.6)
BPE in percentage (median, IQR) (20%)	50.4 (32.5)	52.6 (38.8)	49.8 (32.0)
BPE in percentage (median, IQR) (30%)	38.0 (32.4)	38.6 (39.7)	35.3 (31.9)
BPE in percentage (median, IQR) (40%)	29.1 (28.4)	26.9 (36.1)	25.6 (28.3)
BPE in percentage (median, IQR) (50%)	22.7 (24.2)	18.8 (30.4)	18.6 (23.9)

Figure 7.3: Distribution of the amount of fibroglandular tissue (FGT) for women with a *BRCA* mutation and women without a *BRCA* mutation. The box plots show lowest and highest FGT values (outermost horizontal lines), median FGT (central horizontal line), and interquartile range (top and bottom borders of the box) for breast cancer (no/yes).
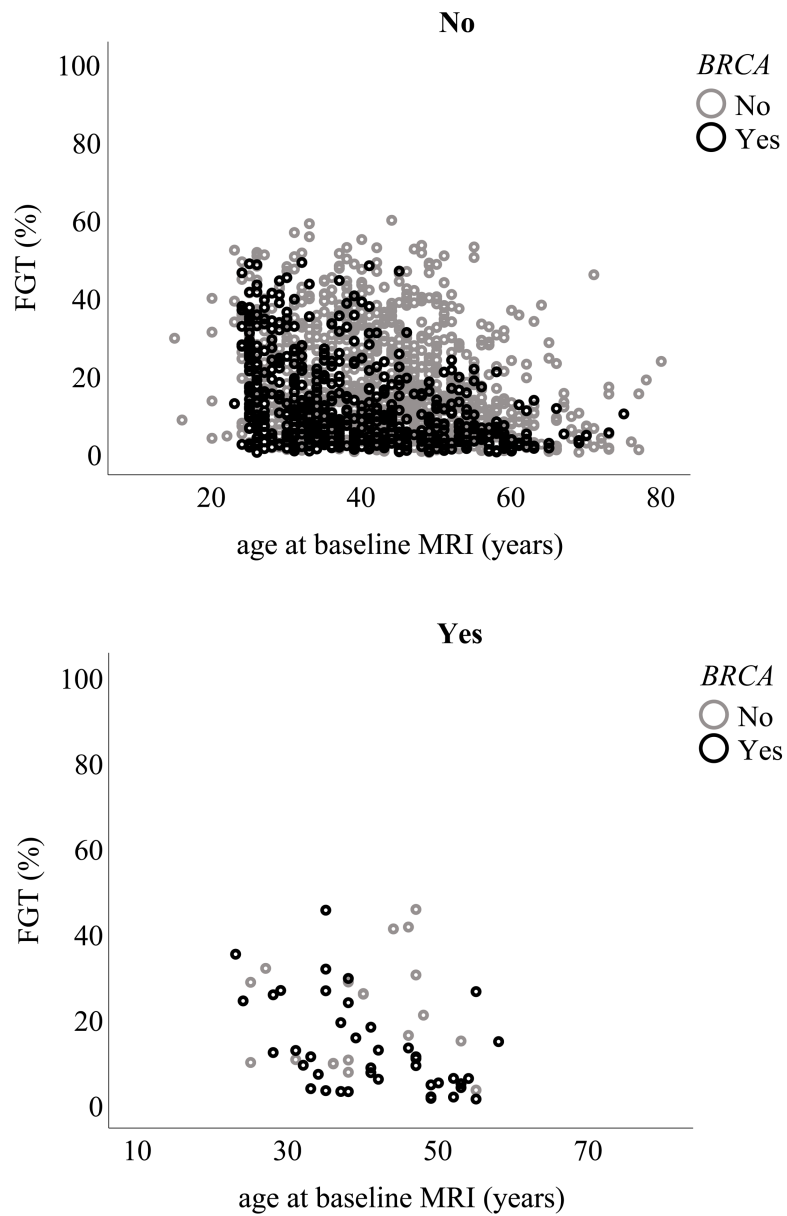
Figure 7.4: Distribution of the amount of fibroglandular tissue (FGT) for women with a *BRCA* mutation and women without a *BRCA* mutation. Scatter plots showing association of FGT to breast cancer occurrence (no/yes) and age at baseline MRI.
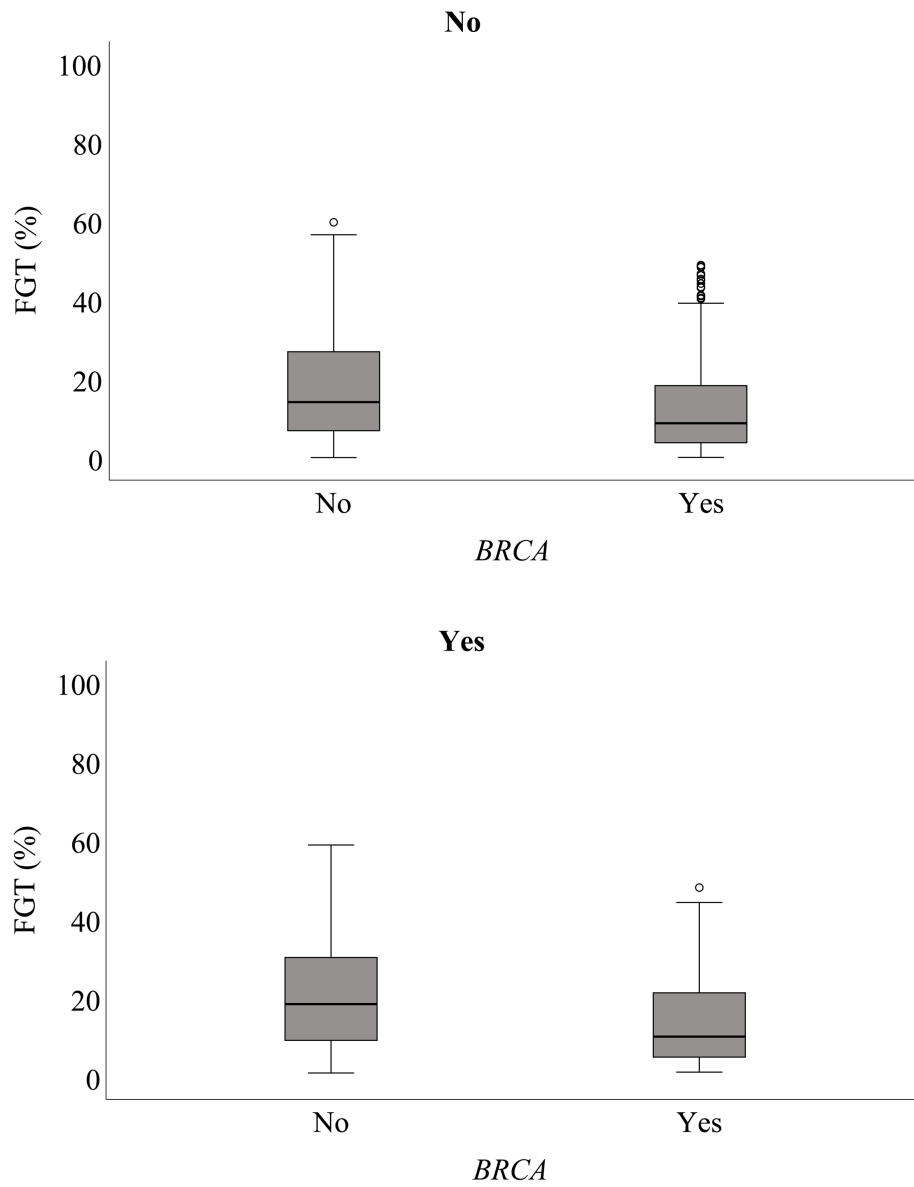
No

Yes

Figure 7.5: Distribution of the amount of fibroglandular tissue (FGT) for women with a *BRCA* mutation and women without a *BRCA* mutation. The boxplots are shown for false positive recall occurrence (no/yes).
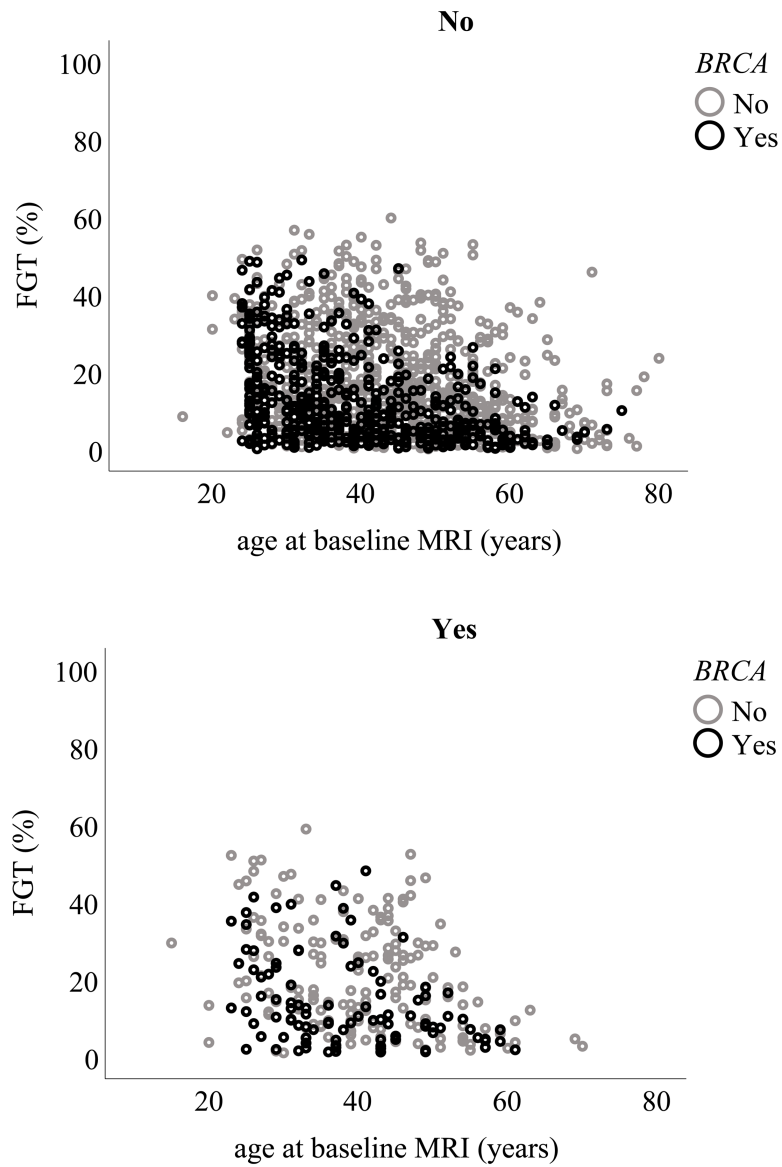
Figure 7.6: Distribution of the amount of fibroglandular tissue (FGT) for women with a *BRCA* mutation and women without a *BRCA* mutation. The scatterplots show the association of FGT to false positive recall occurrence.

No

Yes

Figure 7.7: Distribution of the background parenchymal enhancement (BPE) for women with a *BRCA* mutation and women without a *BRCA* mutation. The box plots show lowest and highest BPE values (outermost horizontal lines), median BPE (central horizontal line), and interquartile range (top and bottom borders of the box) for breast cancer (no/yes).

Figure 7.8: Distribution of the background parenchymal enhancement (BPE) for women with a *BRCA* mutation and women without a *BRCA* mutation. Scatter plots showing association of BPE to breast cancer occurrence (no/yes) and age at baseline MRI.
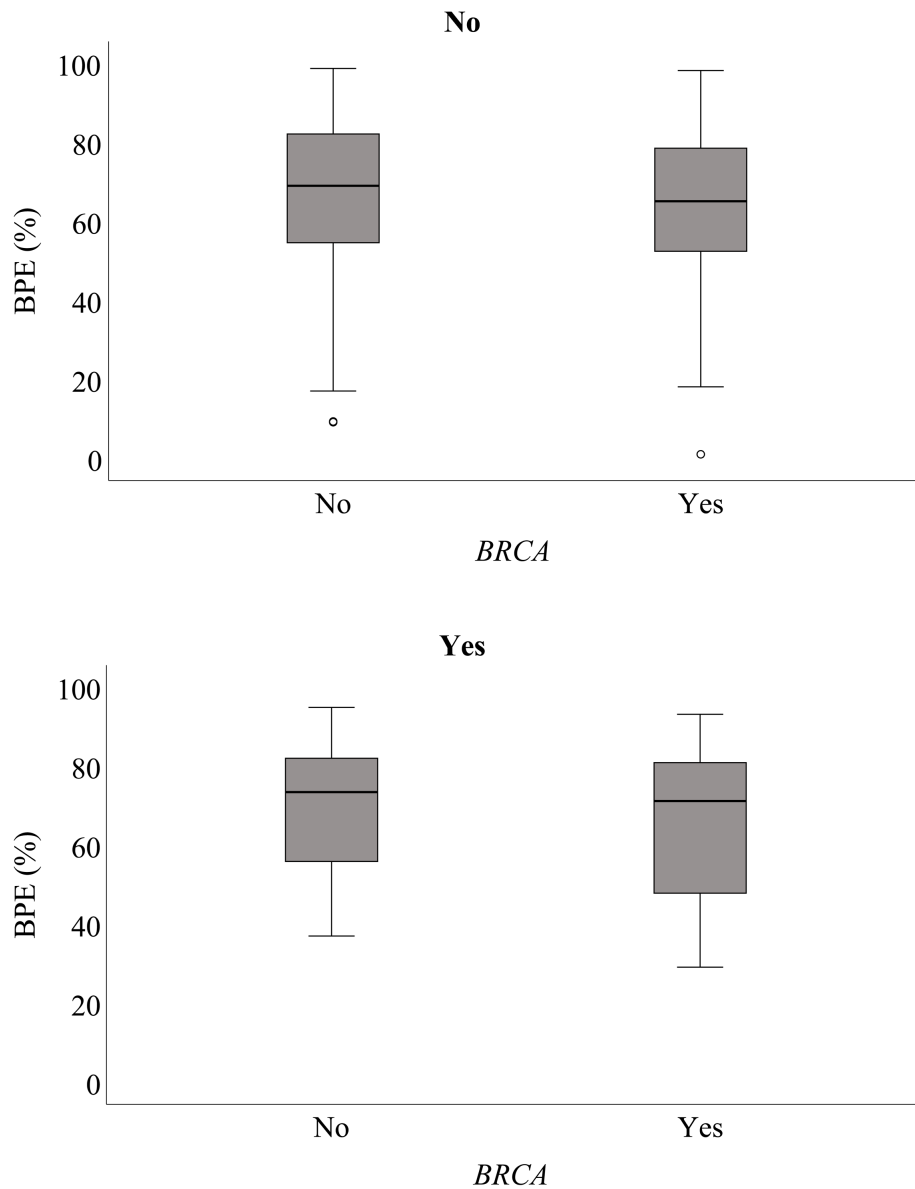
Figure 7.9: Distribution of the background parenchymal enhancement (BPE) for women with a *BRCA* mutation and women without a *BRCA* mutation. The boxplots are shown for false positive recall occurrence (no/yes).
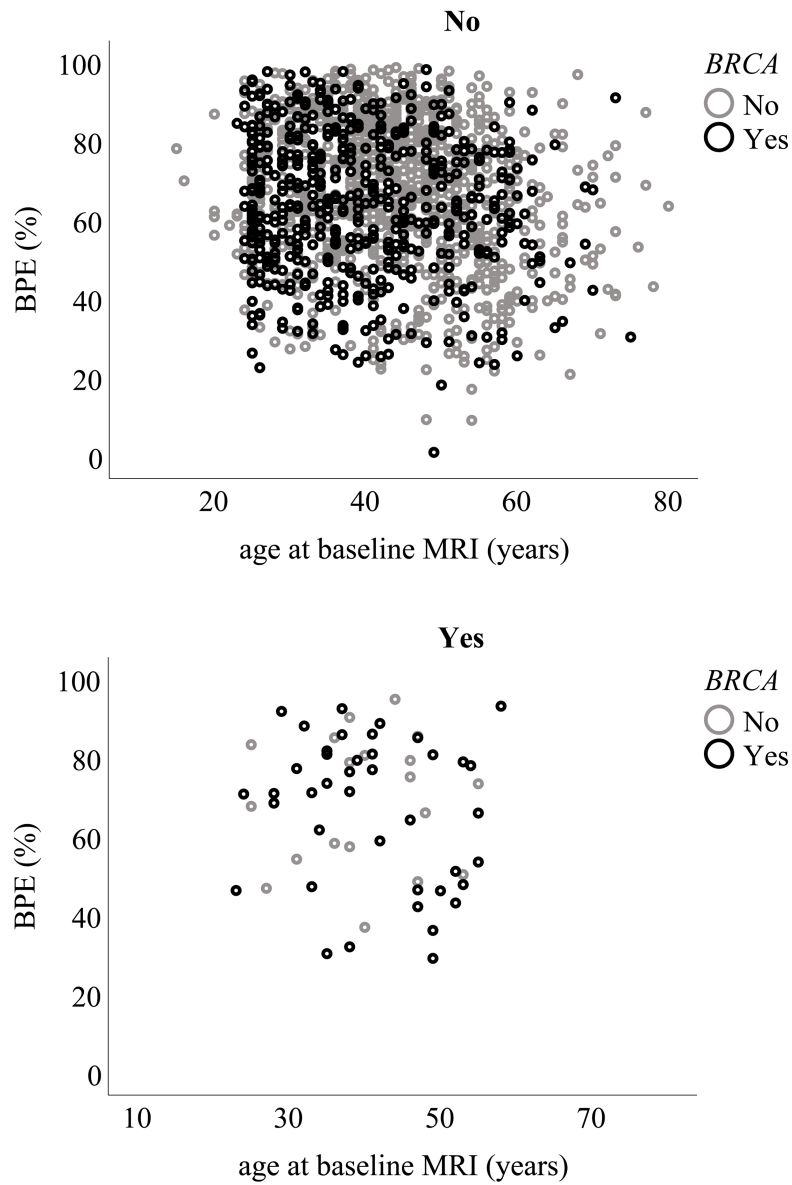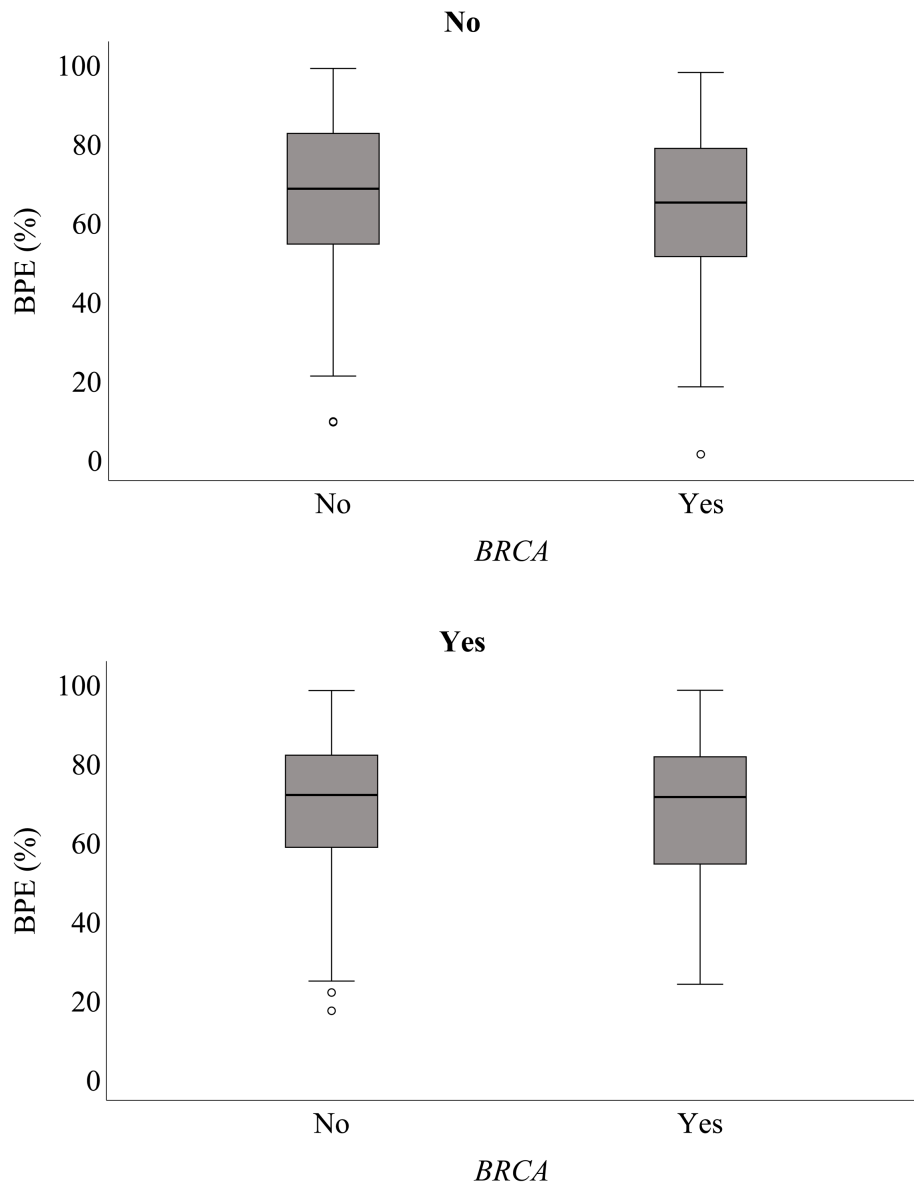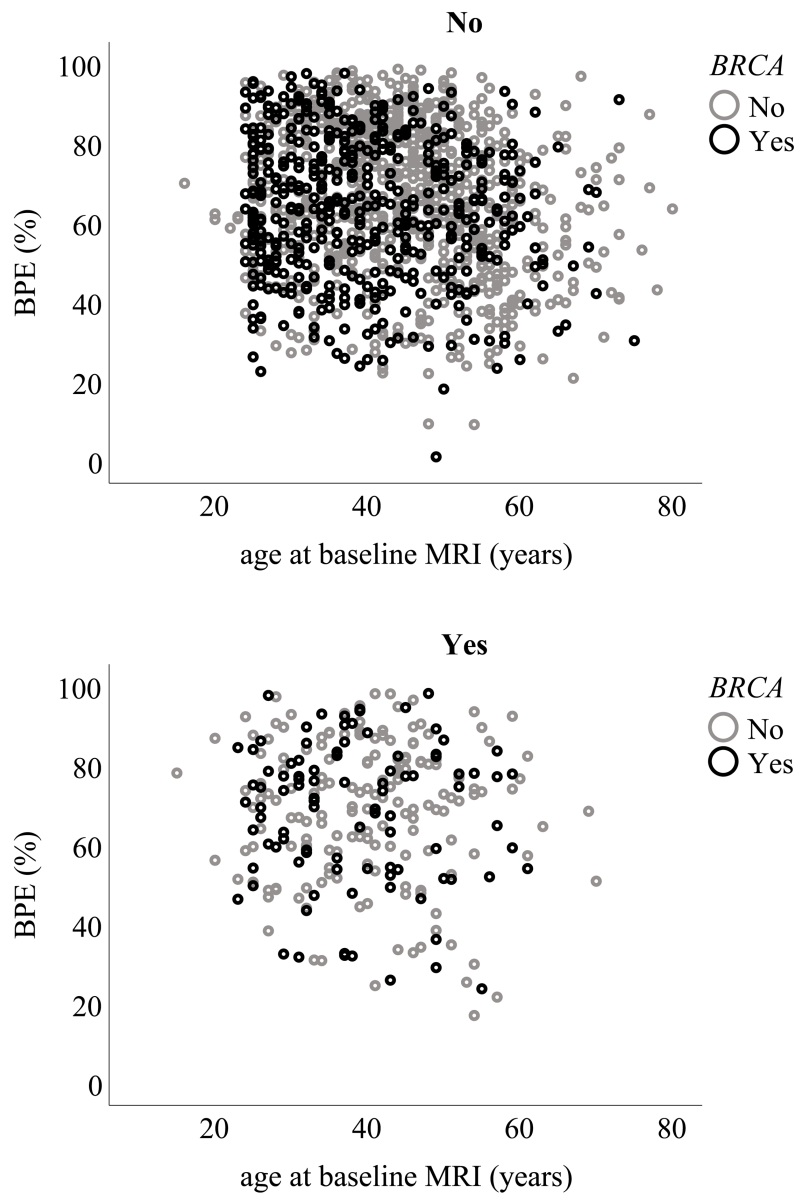
Figure 7.10: Distribution of the background parenchymal enhancement (BPE) for women with a *BRCA* mutation and women without a *BRCA* mutation. The scatterplots show the association of BPE to false positive recall occurrence.

Table 7.7: Regression coefficients and odds ratios for the different BPE cut-offs (20%, 30%, 40%, 50%) for the effect on current and subsequent MRI scans on cancer false positive findings

Model	Predictor	P-value	Included in final model	β (95% CI)	OR (95% CI)	Shrinkage factor
Prognostic cancer model						
Cancer-BPE	*BRCA*	0.001	x	1.285 (0.747 - 1.908)	3.616 (2.110 - 6.743)	0.96
20%-50%	Age	0.930	-			
	BPE*	$\leq$ 0.761	-			
Current false positive recall model						
FPR-BPE	*BRCA*	0.623	-			
20%	Age	0.120	-			
	BPE*	0.005	x	0.404 (0.133 - 0.704)	1.498 (1.142 - 2.022)	0.88
FPR-BPE	*BRCA*	0.605	-			
30%	Age	0.144	-			
	BPE*	0.005	x	0.379 (0.098 - 0.665)	1.461 (1.105 - 1.944)	0.87
FPR-BPE	*BRCA*	0.299	-			
40%	Age	0.046	x	-0.009 (-0.018 - 0.000)	0.991 (0.982 - 1.000)	0.70
	BPE*	0.299	-			
FPR-BPE	*BRCA*	0.299	-			
50%	Age	0.045	x	-0.009 (-0.018 - -0.001)	0.991 (0.982 - 0.999)	0.70
	BPE*	0.424	-			
Prognostic false positive recall model						
FPR-BPE	*BRCA*	0.773	-			
20%-50%	Age	0.001	x	-0.047 (-0.070 - 0.027)	0.954 (0.932 - 0.973)	0.95
	BPE*	$\leq$ 0.839	-			

For every model different shrinkage factors were used, shrunk β and OR are presented. β: standardized coefficients, OR: odds ratio, CI: confidence interval, FGT: amount of fibroglandular tissue, BPE: background parenchymal enhancement

* BPE = low is reference category

Ultrafast and abbreviated breast MRI

8

Ritse M. Mann, Jan C.M. van Zelst, Suzan Vreemann, Roel D.M. Mus

Abstract

Purpose of review Breast MRI has been recognized as the most sensitive modality for breast cancer screening. Its use is, however, restricted due to the high costs of the MRI scan itself, the limited availability of MRI scanners and the long reading times. In this review, the use of shorter MRI protocols for screening is discussed.

Recent findings Shortened scan-protocols have been proposed to reduce both the actual patient handling time and the time to evaluate the scan. So far, these 'abbreviated' scan protocols have shown similar sensitivity and only marginally lower specificity as the common scan protocols, while largely reducing scan time and time required for evaluation. To retain the dynamic information, ultrafast dynamic breast MRI was introduced, capturing the inflow of contrast in a lesion. This comes for free in terms of scan-time. The diagnostic information from this dynamic evaluation is even stronger than the conventional curve types and can be exceptionally helpful in the evaluation of small lesions. Since abbreviated and ultrafast MRI are not mutually exclusive, both techniques can be combined, yielding a screening protocol with a scan-time below five minutes. However, only very few studies evaluated these protocols in a true screening setting, and the number of detected cancers is likewise very low.

Summary Abbreviated protocols with ultrafast MRI allow shortening of breast MRI for screening while retaining excellent sensitivity and specificity. Prospective study data are, however, limited and must be further substantiated.

8.1   Introduction

Worldwide the incidence of breast cancer is growing. Depending on the country, the risk of an average woman to develop breast cancer is up to 15%[1]. Treatment options are simpler, less mutilating and better tolerated when cancers are detected at an early stage; moreover, early detection is still very important for eventual disease-free and overall survival, despite the improvement in systemic therapies[215]. Consequently, screening for breast cancer is employed worldwide, mainly using mammography, although in some Asian countries ultrasound screening is preferred mainly due to the relatively small and dense breasts of women in these countries.

The mammographic screening programs have been highly successful in reducing breast cancer-related mortality, with reported breast cancer-related mortality reductions up to 58% for women who actually attended[216]. Nevertheless, mammographic screening still fails to detect breast cancers early enough for curative treatment in up to 30% of women. These women eventually develop distant metastasis and may die from the disease. Mammography functions especially poor in women at increased risk for the development of breast cancer, with a reported sensitivity well below 50%[14,37]. After the discovery of the high sensitivity of contrast-enhanced breast MRI in the mid-eighties[35,217], it was soon recognized that breast MRI may have a considerable advantage over mammography for early breast cancer detection. Studies focused on the detection of breast cancer in women at increased risk, not only because the performance of mammography in these women was poor, but also due to the fact that the higher incidence of breast cancer in these women required fewer patients to be included in the trials to yield significant results. The results of these initial breast MRI screening trials in many countries were unequivocal, despite the variation in patient populations; the sensitivity of breast MRI is roughly double that of mammography[14,37,41,75,102]. In addition, breast MRI triples the additional yield of breast US as a supplementary screening test for women at increased risk[133]. Finally, the use of breast MRI as a screening tool induces a stage shift in these women and a substantially lower risk for distant metastasis and death (0 - 16% and 0 - 7%, respectively), when compared to historical cohorts (27 - 30% and 17 - 25%, respectively)[39,218].

Based upon the initial MRI screening studies, the American Cancer Society published its recommendation to screen women with a lifetime risk of 20 - 25% or higher annually with breast MRI in 2007. This advice was subsequently also adopted by the European Society for Breast Imaging (EUSOBI), as it was mainly based upon European trials[11,12]. Nonetheless, many national organizations do not adhere to this

advice due to issues with cost-effectiveness. Whereas in women with a BRCA mutation MRI screening is undoubtedly cost-effective[219], this is questioned for women at lower risk, due to the higher amount of women that need to be scanned to find one additional cancer. Unfortunately, the cost-effectiveness is most influenced by the price of a single MRI scan[220].

Meanwhile, further evaluation of MRI as a screening tool in populations at lower risk is ongoing. In a recently published study with very liberal inclusion criteria the supplemental cancer detection using breast MRI in women at average risk, with a negative screening mammogram was 15.5 per thousand in the first round, which reduced to 6.9 per thousand in incidence rounds (nearly all only detected with MRI), fairly equal to the mammographic detection rates in Europe, and, therefore, once again showing the potential of breast MRI to induce a stage shift in breast cancer detection[221]. Further prospective randomized trials in women with extremely dense breasts and women with an increased risk based upon family history alone are ongoing[9,222]. Despite the very likely highly positive results of these trials in terms of early cancer detection and reduction of interval cancers, it can be anticipated that cost-effectiveness analysis will once again show that breast MRI in these populations is in general not affordable, thus limiting the use of an otherwise highly effective screening tool.

For the further implementation of breast MRI as a screening tool it is thus mandatory to reduce the costs of the examination itself, which can be most easily achieved by increasing the throughput of patients in the MRI scanner.

8.2   Abbreviated breast MRI

Abbreviated MRI aims to reduce breast MRI to a minimum length in both acquisition and interpretation time. Nonetheless, most research focuses on reducing the acquisition time of sometimes lengthy breast MRI protocols (reported scan-times of up to 45 minutes). It is, therefore, necessary to define a maximum scan-time for an MRI protocol to be called "abbreviated". In this paper, we will assume that a truly abbreviated breast MRI protocol has a scan-time of less than 5 minutes. Since patient handling times vary between 5 and 10 minutes, this allows the performance of between 4 and 6 breast MRI scans per hour on an average MRI scanner. Virtually all abbreviated protocols use a subset of the sequences employed in current state-of-art multi-

parametric protocols that consist usually of at least a dynamic contrast-enhanced series of about 7 minutes, a T2-weighted acquisition, and possibly a diffusion-weighted series. Such a protocol lasts between 15 and 20 minutes, lengthier protocols are outside of research situations usually not needed and discouraged[11,44]. The studies on abbreviated MRI so far usually acquired a whole multi-parametric protocol and analyzed subsets of sequences blinded to the rest of the protocol. There are hardly any studies that used abbreviated MRI alone for screening of women.

Kuhl et al., who also coined the term, presented the simplest form of abbreviated MRI. In their protocol, a standard T1-weighted sequence was performed before and after administration of a contrast agent[76]. This protocol takes 3 minutes of scan time as each of the T1-weighted acquisitions lasts approximately 60 seconds and a 60-second pause from the start of contrast injection is observed to allow the contrast agent to flow into eventual lesions. Image evaluation was performed by generating maximum intensity projections (MIP) from the post-contrast subtraction images (see Figure 8.1). Reading these is extremely fast (reported to be below 2 seconds), and allows quick dismissal of normal cases when the quality of the MIP images is high. Using this approach an overall sensitivity of 91% using only the MIP images and 100% when also evaluating the complete subtraction series was achieved (compared to the normal multiparametric protocol). The specificity was 94.3%, with a negative predictive value of 99.8%. The main advantage of the full diagnostic protocol was a better characterization of lesions classified as probably benign (i.e. BI-RADS 3, which constituted 9% of abbreviated protocol readings). While the entire study included 606 screens, it only detected 11 cancers, therefore solid conclusions about sensitivity and specificity cannot be drawn. In a study by Mango et al. evaluating cancers only the sensitivity of evaluating the MIP images alone was also substantially lower than when evaluating the actual subtraction images (84 - 96%, vs 93 - 98%, over different readers)[52], and it must be stressed that dismissing studies as normal based upon the MIP images alone requires very good quality MIP images without motion artefacts. Subsequent studies evaluated different variations of abbreviated protocols. Results are listed in Table 8.1. Of note is that several abbreviated protocols do not conform to the definition of a true abbreviated protocol given above, as the scan time is far longer than would be acceptable for high throughput screening.

Almost all authors reported that sensitivity of abbreviated MRI is virtually equal to the sensitivity of a full diagnostic protocol. However, results on specificity are more heterogeneous. While several authors did not report any deviation, many authors report a small to substantial decrease in specificity when evaluating abbreviated breast
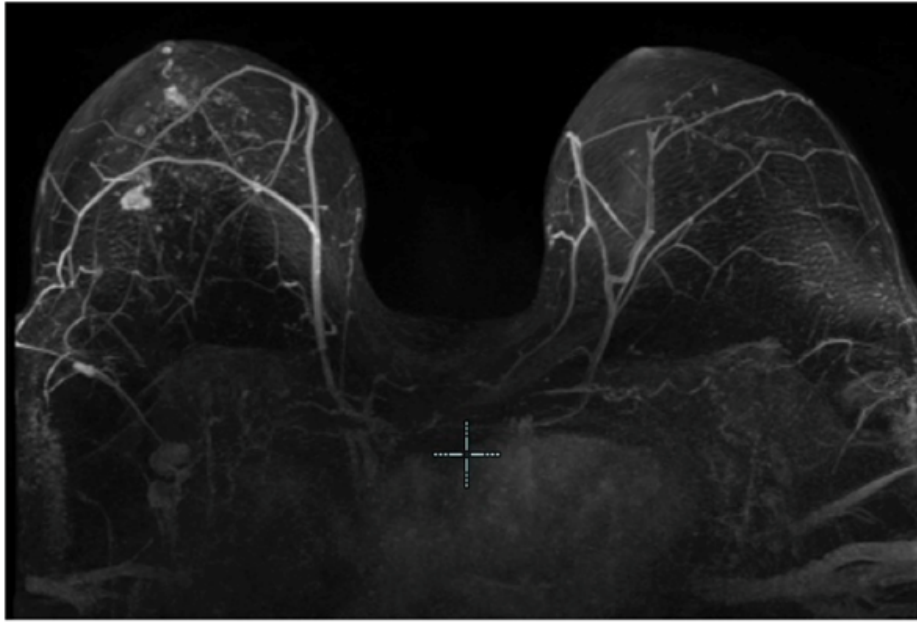
Figure 8.1: Maximum intensity projection of subtracted post contrast acquisition, clearly showing a multifocal invasive lobular cancer in the left breast

MRI[16,17,223–225]. This is in fact also in line with the initial study of Kuhl et al. and seems to be related to the evaluation of the more difficult "probably benign" and relatively small lesions. Heacock et al. reported that lesion conspicuity in malignant breast lesions could be increased by adding a T2-weighted sequence to the abbreviated protocol[54], whereas Chen et al. improved specificity substantially by adding a diffusion-weighted sequence[224]. Several authors reduce the number of post-contrast T1-weighted acquisitions to more than one, thus still enabling dynamic evaluation of breast lesions in the abbreviated protocol, which is likewise aimed at preserving the specificity as it does not seem to have an effect on sensitivity. Recently Strahle et al. tried to define an optimal abbreviated protocol by assessing the diagnostic value of each of the sequences. They obtained the best results by reducing the full protocol to a shortened, but still dynamic T1-weighted series, interleaved with a normal T2-weighted sequence for a total acquisition time of 7.5 minutes, underlining the added discriminatory value of dynamics and to lesser extent T2[226]. It should be noted that their full protocol, however, did not include DWI. In conclusion, it is evident that abbreviated MRI allows shortening of the full diagnostic breast MRI protocol, without a significant reduction in sensitivity. Effects on specificity are mainly limited to patients with lesions of low suspicion for malignancy and can be improved by either preserving dynamic evaluation, or adding T2 and DWI. However, whether the extension in acquisition time required for improved lesion classification is economically viable in a situation where most scans are completely negative (i.e. in screening,

especially non-high-risk) remains debatable.

8.3 Ultrafast breast MRI

Ultrafast breast MRI is developed to capture the inflow of contrast in breast lesions and hence enable dynamic analysis of contrast wash-in rather than contrast wash-out. The technique builds largely on earlier attempts to analyze the pharmacokinetic characteristics of breast cancer vasculature, which were shown to allow good separation between benign and malignant breast lesions[227,228]. The main difference
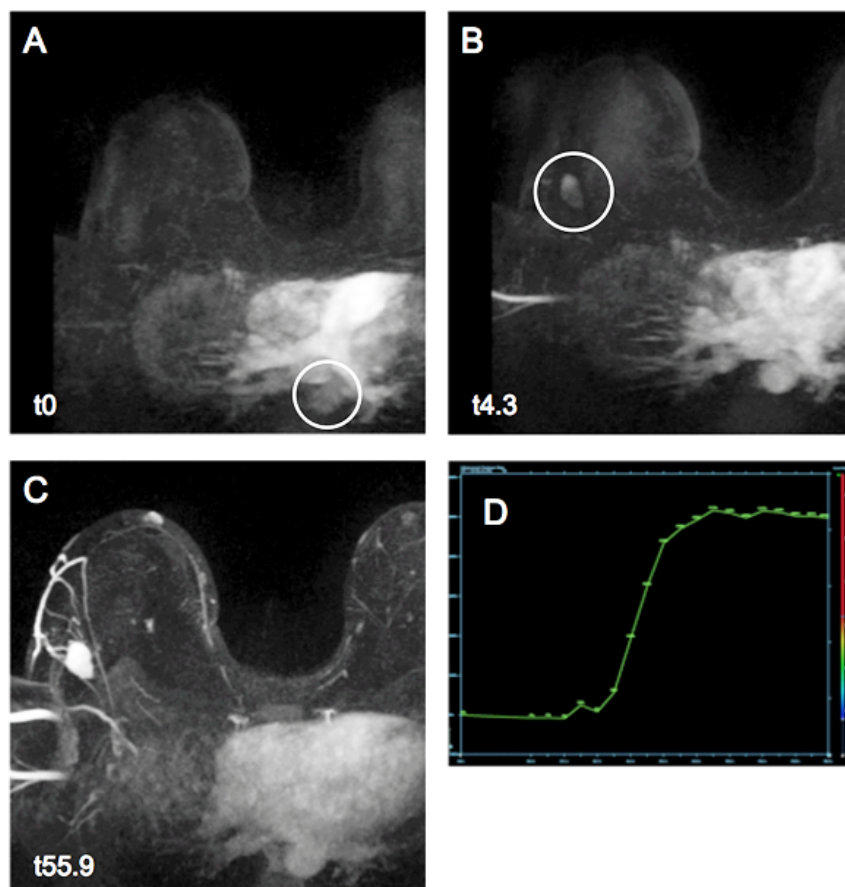
Figure 8.2: Maximum intensity projections from an ultrafast dynamic series with a temporal resolution of 4.3 seconds. TTE is relative to the enhancement of the descending aorta, the first volume that shows this vessel is designated t0 (A). In (B) the first volume where the lesion is visible is shown, the very early enhancement (only 4.3 seconds after the aorta) make it very likely that this lesion is malignant. In later time points, morphology becomes clear and can be easily appreciated (C). In (D) the time versus signal intensity curve for the inflow period is presented. The very steep upslope is indicative of malignancy.

between these techniques, however, is that ultrafast breast MRI is performed with a spatial resolution that enables morphological assessment of eventual breast lesions as it meets the international standards for diagnostic breast MRI, and can therefore also be used for lesion detection. Like abbreviated MRI the term "ultrafast" is not clearly defined and sometimes used for MRI techniques that are not exceptionally fast. In this paper, ultrafast breast MRI is defined as T1-weighted imaging with a temporal resolution below 6 seconds, and a spatial resolution of at least 1x1x2.5 mm.

Herrmann et al. were in 2011 the first to publish that, view-sharing, a then novel approach to k-space filling, allowed construction of high-resolution breast MRI acquisitions at a temporal resolution of 5.7 seconds[78]. The commercial name of the used sequence is time-resolved angiography with stochastic trajectories (TWIST), and hence TWIST has become sort of synonymous with ultrafast breast MRI. Le et al. and Saranathan et al. subsequently showed that similar results were also possible using fat-saturated and even Dixon techniques, although these are (still) computationally heavy and, therefore, require long reconstruction times[229,230].

In 2014, we proposed the use of ultrafast breast MRI alone as a new screening technique for breast cancer, based upon a study of 199 abnormalities where the wash-in characteristics obtained with TWIST were compared to the wash-out characteristics from the same lesions obtained from regular T1-weighted series. The ultrafast protocol obtained 20 T1-weighted volumes at a temporal resolution of 4.3 seconds, for a total duration of 102 seconds, therefore being faster than any of the abbreviated protocols discussed above. The study unequivocally showed that maximum slope (MS), a heuristic dynamic parameter obtained from the TWIST images largely outperformed the conventional curve types (area under the curve 0.81 vs 0.69), while still allowing imaging at diagnostic spatial resolution. MS can be assessed in a similar way as the curve types, where a steep slope is indicative of malignancy, a shallow slope of a likely benign lesion and an intermediate curve is indeterminate[55]. In further studies, it was subsequently shown that also the time to enhancement relative to aortic enhancement (TTE) is a highly discriminative and very reproducible parameter for breast lesion characterization[231]. In simple words, lesions that enhance within 10 seconds after the aorta are likely malignant, lesions that enhance between 10 and 15 seconds after the aorta are indeterminate, and lesions that enhance more than 15 seconds after the aorta are likely benign. As ultrafast breast MRI allows the generation of a series of MIP images (one for each post-contrast timepoint acquired), it is possible to evaluate these dynamic MIPS as a movie of contrast inflow. Based upon the above, the first enhancing lesion is usually the one that deserves the most

attention. In patients with cancer, you may observe a "lightbulb" effect, as you see the cancer enhancing in an otherwise completely black breast (see Figure 8.2). This effect is also very valuable for the evaluation of breasts with nodular background enhancement, to differentiate true lesions from focal areas of adenosis[77]. Another approach is the assessment of the time between arterial supply and venous drainage in breast lesions. Overall, when arterial supply is observed, the lesion already has a very high likelihood of being malignant, which further increases when the time from arterial supply to venous drainage is short, a parameter that is highly related to the TTE described above[232]. It is, nonetheless, suggested to start with the evaluation of the MIP generated from the last time point obtained. When this MIP is of good quality and truly negative there is no need for evaluation of the other phases (which is comparable to the MIP reading proposed by Kuhl et al. and Mango et al.). Of note is that the excellent discriminating capacity of MS and TTE are also valid for computer-aided diagnosis[233], and remain valid over various types of breast cancer, including invasive lobular cancers and ductal carcinoma in situ[234]. It is still essential to realize that, even though with TWIST a spatial resolution is obtained of 1x1x2.5 mm, this is still lower than in most high-resolution T1-weighted protocols where voxels usually are below 1mm isotropic. Consequently, multi-plane reconstruction is somewhat limited and therefore morphological evaluation is not as good as with conventional protocols in all directions. However, recent studies employing compressed sensing with iterative reconstruction rather than view-sharing allow a substantial further reduction of the spatial resolution, down to 0.8x0.8x1.6 mm, in the same time frame[235,236].

Whether ultrafast MRI alone is good enough for breast screening remains to be evaluated, as published studies so far mainly looked at lesion classification capacity. However, Van Zelst et al. presented at RSNA 2016 preliminary results of a reader study with TWIST only versus a full diagnostic protocol on 85 biopsied lesions, including 31 malignancies, obtained from over 3000 screening examinations, and complemented with 115 normal cases. Based upon the first 4 readers the average sensitivity of TWIST only evaluation was 80%, compared to 85% for the full diagnostic protocol, without statistical significance. However, the relative low sensitivities show the difficulty of the comprised case set consisting of only screening cases. Typically the average specificity of TWIST only reading was, at 81%, significantly higher than with the full diagnostic protocol at 77%, while the average area under the curve for the two reading modes was virtually identical[81]. Nevertheless, the final results of that study, including more readers must be awaited.

8.4   Hybrid imaging

It is important to realize that the definitions of abbreviated and ultrafast MRI are not mutually exclusive. In fact, ultrafast MRI alone is a form of an abbreviated protocol with a very short acquisition time (less than 2 minutes). However, for the performance of ultrafast breast MRI as a stand-alone technique, specific sequences are required that are not uniformly available. With classic, and far more commonly available, keyhole techniques (i.e. turboFLASH, eThrive etc.), it is, unfortunately, not possible to obtain the required temporal resolution at a diagnostically acceptable spatial resolution. Nevertheless, even in the most basic form of an abbreviated protocol, the period of contrast inflow must be observed. Kuhl et al. solved this by simply waiting 60 seconds before acquiring the post-contrast acquisition. Obviously, ultrafast acquisitions (even of lower spatial resolution) can be inserted in this waiting period without any penalty in overall scanning time. This still allows very fast evaluation of the entire protocol by evaluating the MIP image and high-resolution subtraction series, but in the presence of lesions substantial additional dynamic information is present for lesion classification (it should be noted that when the last post-contrast scan is negative, it is not needed to evaluate all the previous phases). The superior performance of wash-in characteristics from ultrafast imaging compared to standard wash-out evaluation truly makes late phase acquisitions obsolete and enables a huge shortening of the protocol without losing any information. Whether or not the addition of T2 or DWI is beneficial in the screening setting remains to be seen.

8.5   Conclusion

The further expansion of breast MRI as a screening tool is limited by the high cost of the examination itself. To enable a more widespread use of this highly effective screening technique, it is important to reduce the price of an actual MRI scan. The most practical approach to that is shortening the actual acquisition protocol, as only an increase in throughput can reduce the price of a single breast MRI to a more acceptable level. This is particularly important when screening women at an intermediate risk for the development of breast cancer, such as those with a positive family history without hereditary genetic mutations, and women with extremely dense breasts. Abbreviated protocols, particularly those employing (also) ultrafast breast MRI, may offer the required efficiency. However, although in recent years many studies have investigated such protocols, the actual amount of screening-detected cancers in these studies is still very low. This makes a true assessment of the value of either abbreviated or ultrafast MRI as a screening technique still somewhat lim-

Table 8.1: Studies comparing abbreviated protocols to full diagnostic protocols; effects on sensitivity and specificity

Author (year)	Abbreviated protocol (duration)	Number of cancers/total cases	Sensitivity Abbreviated protocol	Specificity Abbreviated protocol	Sensitivity Full diagnostic protocol	Specificity Full diagnostic protocol
Kuhl 2014[76]	T1 pre & post contrast (3 min)	11/606 (screening population)	100%	94.3%	100%	93.4%
Mango 2015[52]	Sag T1 pre & post contrast (10-15 min)*	100 (cancers only)	93-98%	NA	95-98%	NA
Grimm 2015[a][53]	T2 FS, T1 pre & post contrast (21 min)*	12/48 (enriched)	86%	52%	95%	52%
Grimm 2015[a][53]	T2 FS, T1 pre & post contrast + second post contrast (23 min)*	12/48 (enriched)	89%	45%	95%	52%
Harvey 2016[16]	T1 pre & post contrast (4.4 min)	7/568 (screening population)	100%	94%	100%	96%
Heacock 2016[a][54]	T1 pre & post contrast (7min)	107 (cancers only)	97.8%	NA	99.4%	NA
Heacock 2016[a][54]	T2 FS, T1 pre & post contrast (12 min)	107 (cancers only)	99.4%	NA	99.4%	NA
Moschetta 2016[17]	T2 STIR, T2, T1 pre & postcontrast (10 min)	75/470 (clinical population)	89%	91%	92%	92%
Chen 2017[223]	T1 pre & post contrast (3 min)	16/478 (dense breast screening)	93.8%	88.3%	100%	94.6%
Petrillo 2017[237]	T1 pre & post contrast (3 min)	207/508 (clinical population)	99.5%	75.4%	99.5%	77.1%
Chen 2017[a][224]	T1 pre & post contrast (3 min)	14/356 (dense breast screening)	92.9%	86.5%	100%	96.8%
Chen 2017[a][224]	T1 pre & post contrast +DWI (6 min)	14/356 (dense breast screening)	100%	95%	100%	96.8%
Romeo 2017[225]	T1 pre & 3 * post contrast (7 min)	110/180 (lesions only)	99%	93%	97%	95%
Panigrahi 2017[238]	T1 pre & post contrast (3 min)	11/678 (high risk screening)	81.8%	97.2%	81.8%	97.4%

* scan time reported included patient positioning
[a] reported 2 different abbreviated protocols in one study

ited. Considering the current results, it is nevertheless unlikely that the sensitivity of abbreviated MRI is much lower than that of a full diagnostic protocol and defini-

tively much higher than either mammography, tomosynthesis, or ultrasound. This already underlines its potential value for supplemental evaluation of women with very dense breasts, even when it is still somewhat unsure whether it can completely replace the full diagnostic protocol in women currently already screened with MRI. It is, in fact, likely that the specificity of abbreviated breast MRI is somewhat lower than that of a full diagnostic protocol, which in the screening setting might be a large issue, leading to a substantially higher number of recalls and additional evaluations. Still, this may already be largely compensated by incorporation of ultrafast breast MRI in the abbreviated protocol (if not used as stand-alone technique). Consequently, ultrafast abbreviated MRI will likely become the MRI screening technique of the near future.

Compressed Sensing for breast MRI

9

Suzan Vreemann, Alejandro Rodriguez-Ruiz; Dominik Nickel, Laura Heacock, Linda Appelman, Jan C.M. van Zelst, Nico Karssemeijer, Elisabeth Weiland, Marnix C. Maas, Linda Moy, Berthold Kiefer, Ritse M. Mann

Abstract

Objective Ultrafast dynamic contrast-enhanced magnetic resonance imaging of the breast enables assessment of the contrast inflow dynamics while providing images with diagnostic spatial resolution. However, the slice thickness of common ultrafast techniques still prevents multiplanar reconstruction. In addition, some temporal blurring of the enhancement characteristics occurs in case view-sharing is used. We evaluate a prototype compressed-sensing volume-interpolated breath-hold examination (CS-VIBE) sequence for ultrafast breast MRI that improves through plane spatial resolution and avoids temporal blurring while maintaining an ultrafast temporal resolution (less than 5 seconds per volume). Image quality (IQ) of the new sequence is compared with an ultrafast view-sharing sequence (time-resolved angiography with interleaved stochastic trajectories [TWIST]), and assessment of lesion morphology is compared with a regular T1-weighted 3D Dixon sequence (VIBE-DIXON) with an acquisition time of 91 seconds.

Materials and Methods From April 2016 to October 2016, 30 women were scanned with the CS-VIBE sequence, replacing the routine ultrafast TWIST sequence in a hybrid breast MRI protocol. The need for informed consent was waived. All MRI scans were performed on a 3T MAGNETOM Skyra system (Siemens Healthcare, Erlangen, Germany) using a 16-channel bilateral breast coil. Two reader studies were conducted involving 5 readers. In the first study, overall IQ of CS-VIBE and TWIST in the axial plane was independently rated for 23 women for whom prior MRI examinations with TWIST were available. In addition, the presence of several types of artefacts was rated on a 5-point scale. The second study was conducted in women (N = 16) with lesions. In total, characteristics of 31 lesions (5 malignant and 26 benign) were described independently for CS-VIBE and VIBE-DIXON, according to the BI-RADS MRI lexicon. In addition, a lesion conspicuity score was given.

Results Using CS-VIBE, a much higher through-plane spatial resolution was achieved in the same acquisition time as with TWIST, without affecting in-plane IQ (P = 0.260). Time-resolved angiography with interleaved stochastic trajectories showed slightly more motion artefacts and infolding and ghosting artefacts compared with CS-VIBE, whereas CS-VIBE showed more breathing and pulsation artefacts. For morphologic assessment, intrareader agreement between CS-VIBE and the more time-consuming VIBE-DIXON was slight to almost perfect and generally higher than interreader agreement. Mean sensitivity (84.0% and 92.0% for CS-VIBE and VIBE-DIXON, P = 0.500) and specificity (60.0% and 55.4% for CS-VIBE and VIBE-DIXON, P = 0.327) were comparable for both sequences.

Conclusions Compressed-sensing volume-interpolated breath-hold examination al-

lows an increase of the through-plane spatial resolution of ultrafast dynamic contrast-enhanced magnetic resonance imaging compared with TWIST at a comparable in-plane IQ. Morphological assessment of lesions using CS-VIBE is comparable to VIBE-DIXON, which takes 18 times longer. Consequently, CS-VIBE enables 3D evaluation of breast lesions in ultrafast breast MRI.

9.1   Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) of the breasts is considered to be the most sensitive method for early breast cancer detection. In a meta-analysis, an overall sensitivity of 90% was reported, and also in recent screening studies in high-risk patients, a sensitivity of over 90% was achieved[41,75,239]. Nonetheless, breast MRI is currently only used for screening purposes in women at increased risk for developing breast cancer ($> 20\%$ of lifetime risk)[166]. In the early years of breast MRI, a lower specificity of DCE-MRI in comparison to mammography[240] was repeatedly reported. This led to the development of dedicated protocols for breast MRI that incorporated, for example, contrast-wash-out evaluation, late high-resolution T1 acquisitions, T2, and diffusion sequences. Although several authors showed that each of these techniques has the ability to improve the specificity of breast MRI without lowering its sensitivity[57,158,241], their added value in a screening setting where most scans are normal is debatable. In fact, screening women at lower risk is only cost-effective if the costs of breast MRI itself can be reduced, and thus any sequence that lengthens the protocol should be critically evaluated.

Current research hence focuses on shortening breast MRI protocols without decreasing sensitivity or specificity. For this purpose, abbreviated protocols were created and investigated[52–54,76]. In the most basic form, an abbreviated MRI protocol consists of T1 acquisitions before and after contrast administration (typically around 90 seconds per acquisition), subtraction, and the formation of maximum intensity projections. This allows morphologic evaluation of possible breast lesions but discards most other information, including the dynamic contrast characteristics that have long been regarded as a hallmark of breast MRI.

Ultrafast MRI protocols re-enable dynamic analysis during contrast inflow and therefore do not lengthen the protocol (temporal resolution of 5 seconds). In fact, ultrafast breast MRI can be performed within 2 minutes and thus even shortens the protocol further. Mann et al[55] showed that the dynamic parameters obtained can be used to characterize lesions as benign or malignant, which is particularly useful in the evaluation of small breast lesions typically found in screening, as these often do not show the classical morphologic features of malignancy[55].

Although the ultrafast sequences used today can achieve a diagnostic spatial resolution (under 1 x 1 x 2.5 mm) according to international standards, the through-plane resolution is not sufficient for multiplanar reconstruction. In addition, the view-

sharing used may give rise to temporal blurring.

With the recent advances in MRI and computational power, a new technique called compressed sensing (CS) has become available. Compressed-sensing accelerates imaging by acquiring a subset of the full k-space. Random subsampling of data and constraining the reconstruction for certain properties allow for obtaining artefact-free images[242]. Using CS, an image with a high spatial resolution can be computed at a temporal resolution of less than 5 seconds.

In this study, we evaluate whether a prototype sequence for ultrafast dynamic breast MRI based on a gradient-echo sequence with compressed sensing (compressed-sensing volume-interpolated breath-hold examination [CS-VIBE])[243] could replace the DCE series currently in use in clinical practice. This requires both a very high temporal resolution to document the wash-in phase and high 3-dimensional spatial resolution to allow for multiplanar assessment of morphology. We compare CS-VIBE to a commercially available technique based on view-sharing (time-resolved angiography with interleaved stochastic trajectories [TWIST]), evaluating image quality (IQ) and the presence of artefacts, to assess whether the compressed sensing approach can overcome the limitations that are still present with the view-sharing technique. In patients with abnormalities, we also compare the CS-VIBE sequence for morphologic evaluation and lesion conspicuity assessment to a regular T1-weighted gradient-echo Dixon sequence (VIBE-DIXON) with an acquisition time of 91 seconds.

9.2   Materials and Methods

Patients

From April 2016 until October 2016, 30 women with a breast MRI indication were scanned with a hybrid protocol that included the prototype ultrafast CS-VIBE sequence during contrast inflow. The institutional review board waived the need for informed consent.

Imaging

All women were scanned on a 3T MAGNETOM Skyra system (Siemens Healthcare, Erlangen, Germany) using a 16-channel bilateral breast coil. In our previously published hybrid protocol,13 the regular ultrafast view-sharing acquisition (TWIST) is interleaved in between the initial phases of a high-resolution T1-weighted 3D gradient echo Dixon sequence (VIBE-DIXON). In this study, the TWIST sequence was replaced with the prototype CS-VIBE sequence[243]. The CS-VIBE sequence incoher-

Table 9.1: Sequence parameters

	CS-VIBE	TWIST	VIBE-DIXON
TR (ms)	4.47	3.96	5.97
TE (ms)	2.06	2.02	1: 2.46 2: 3.69
FA (°)	15	20	15
Field of View (mm^2)	358.4x358.4	345.6x345.6	332.8x332.8
Matrix size	448x381	384x346	416x358
In-plane resolution (mm)	0.8x0.8	0.9x0.9	0.9x0.9
TWIST central region A (%) / Sampling density B (%)	-	A: 15 B: 10	-
Slice thickness (mm)	1.6	2.5	1.0
Voxel volume (mm^3)	1.02	2.03	0.81
Phase resolution (%)	85	90	86
Slice resolution (%)	70	77	80
Acceleration factor	20	3	3
Acceleration mode	CS	GRAPPA	GRAPPA
Time resolution per volume (s)	1.55	4.57	91
Total acquisition time (min)	1:40	2:03	1:31

ently subsamples the phase-encoding plane in a variable-density manner with a selectable overall acceleration factor. Specifically, the sampling is designed to obey a Gaussian distribution for which the density drops from center to border of acquired k-space and can be adjusted in both phase-encoding directions. Furthermore, it is also guaranteed that the accumulated sampling up to any time point does not deviate from its expectation value by more than one. Image reconstruction is performed by optimizing a cost function of the form

$$\min_{\mathbf{I}}(\frac{1}{2}||\mathbf{AI} - \mathbf{D}||_2^2 + ||\mathbf{WI}||_{\mathbf{I}}) \tag{9.1}$$

where $\mathbf{I}$ is the reconstructed 3D + time image series; $\mathbf{D}$ is the acquired MRI k-space data; and $\mathbf{A}$ is the system operator consisting of multiplication with coil sensitivities, Fourier transformation, and masking. Finally, $\mathbf{W}$ is the redundant Haar wavelet transformation that was applied in the phase-encoding directions and the time direction. Parameters of the prototype CS-VIBE, the TWIST, and the VIBE-DIXON are presented in Table 9.1.

In addition, to enable comparison of CS-VIBE to TWIST, we collected all previous MRI acquisitions of these women that contain TWIST from the picture archiving and communication system, when available. Images were obtained before, during, and after contrast administration (0.1 mmol/kg Dotarem [gadoteric acid 0.5 mmol/mL];

Guerbet, Villepinte, France) through an intravenous cannula placed in the cubital vein. A power injector (Medrad, Warrendale, PA) was started immediately after completion of the first population in k-space in case of the TWIST sequence or after the first acquisition in case of the CS-VIBE sequence at a rate of 2.5 mL/s. The contrast bolus was followed by a saline flush of 20 mL.

Reader studies

Two separate reader studies were conducted, using an in-house developed platform for diagnostic studies[244]. This platform showed original and subtracted images for all available time points and provided reformatted images in the coronal and sagittal planes, along with the originally obtained axial acquisitions. Five readers with various levels of experience read both studies (1 radiologist with 9 years of experience in breast MRI, 1 radiologist with 7 years of experience in breast MRI, 1 radiologist with 3 years of experience in breast MRI, 1 radiologist with less than 1 year of experience in breast MRI, and 1 radiologist in training with less than 1 year experience in breast MRI). For both studies, readers were blinded to all clinical information, and the cases were presented to the readers in random order. In the first study, image quality of CS-VIBE was compared with TWIST. In the second study, lesion conspicuity and morphological assessment was compared with VIBE-DIXON.

Reader Study 1: Image Quality of CS-VIBE Compared with TWIST

This study was conducted in 23 women for whom prior MRI examinations with TWIST were available. Overall image quality (IQ) of the CS-VIBE and TWIST images was independently rated on a discrete scale from 1 to 100. In addition, the effect of several types of artefacts (motion, breathing and pulsation, ghosting and infolding, external influences [susceptibility]) on image evaluation was rated on a 5-point scale (1 = heavily disturbing to 5 = not present).

Reader Study 2: Lesion Appearance on CS-VIBE Compared with VIBE-DIXON

The second study was conducted using all lesions described in the clinical evaluation of the breast MRI scans. Motion correction was applied for VIBE-DIXON images[148]. In total, 31 lesions were described in 16 patients. Five lesions were proven to be malignant, 26 lesions were benign (either based upon histology [N = 13] or clinical follow-up of at least 1 year [N = 13]). Patients were presented in random order to the readers. Either the CS-VIBE images or the VIBE-DIXON water images were presented. Each case was thus read twice. Readers were informed about the location of the lesions in the breast(s) (quadrant and clockface); no further information was provided.

The lesions were independently evaluated on the CS-VIBE images and the VIBE-DIXON water images. Each lesion was described according to the BI-RADS MRI lexicon[19,126]. A BI-RADS score (1 to 5) was provided. In addition, a lesion conspicuity score on a 5-point scale (1 = not visible to 5 = very clear) was provided for each lesion.

Dynamic Analysis

To investigate whether the dynamic parameters obtained with the CS-VIBE sequence were in line with those obtained with TWIST, the time to enhancement (TTE) was measured and maximum slope type of the inflow curve was scored by 1 reader for all lesions in which a prior MRI scan with TWIST was available (N = 24; all benign). Time to enhancement was defined as the time to enhancement measured from the moment the aorta starts to enhance[245]. Slope types were scored according to Mann et al[55].

Statistical Analysis

To assess the IQ of CS-VIBE compared with TWIST, we averaged all IQ scores per reader. To account for multiple independent readers and for repeated measures, a linear mixed model approach was used. The sequence (CS-VIBE, TWIST) was used as a fixed classification score. Interreader agreement was assessed using the intraclass correlation coefficient (ICC). To assess the impact of artefacts, we dichotomized the results per reader, based upon the assumption that scores 1 and 2 might impair the evaluation, whereas 3 to 5 are given in cases where the artefacts might be present but will not impact the evaluation. We, subsequently, evaluated whether the presence of severe artefacts (rated 1 or 2) in each of the 4 defined artefact categories was more prevalent in CS-VIBE or TWIST using McNemar tests per-reader. Interreader agreement was assessed using Fleiss kappa for the dichotomized artefact values.

To assess the quality of lesion characterization, we compared lesion conspicuity scores between CS-VIBE and VIBE-DIXON using the Wilcoxon signed ranks test. To assess whether lesion characteristics were comparable for the 2 sequences, we calculated intrareader agreement using weighted kappa statistics. Interreader agreement was assessed using Fleiss kappa statistics. The strength of the agreement was defined as poor for less than 0.000, slight for 0.000 to 0.200, fair for 0.201 to 0.400, moderate for 0.401 to 0.600, substantial for 0.601 to 0.800, and almost perfect for 0.801 to 1.000. Sensitivity and specificity of both sequences based on morphologic evaluation were calculated using the BI-RADS scores, where BI-RADS 3, 4, and 5 were considered

positive. Sensitivities and specificities were calculated using the pooled scores for all readers and compared using McNemar tests. All statistical analyses were performed using SPSS (version 22, SPSS Inc, Chicago, IL). Results with 2-sided P-values of 0.05 or less were considered statistically significant.

9.3 Results

Subjects

The 30 women included in this study were on average 46 years of age (range: 28 - 72). Breast MRI scans were performed for screening in women with a *BRCA* gene mutation (N = 15), family history of breast cancer without a known mutation (N = 3), personal history of breast cancer (N = 3), dense breasts (N = 1), ataxia telangiectasia (N = 1), and previous radiation to the chest (N = 1); problem solving in 1 woman with duct ectasia and 1 woman referred from the screening program; and staging of known breast cancer (N = 4).

Reader Study 1: Image Quality of CS-VIBE Compared with TWIST

Using the CS-VIBE sequence, images could be obtained with a comparable temporal resolution as that of TWIST (see Table 9.1). The advantage of CS-VIBE over TWIST is the possibility of achieving a much higher spatial resolution (0.8 x 0.8 x 1.6 mm vs 0.9 x 0.9 x 2.5 mm) and therefore enabling better multiplanar reconstructions (see Figure 9.1). The average IQ scores per reader were 75, 66, 70, 69, and 56 for CS-VIBE, and 70, 63, 76, 71, and 45 for TWIST (see Table 9.2), and statistical analysis showed that IQ was comparable between CS-VIBE and TWIST (P = 0.260). Table 9.2 also lists the artefact scores per reader. Overall, artefacts that disturbed image interpretation were relatively rare. According to the 5 readers, motion artefacts severely impacted image quality appeared more in TWIST compared with CS-VIBE (0 to 3 cases and 0 to 9 cases for CS-VIBE and TWIST, respectively); in 1 reader (R3, P = 0.031), this difference was significant. Severe breathing and pulsation artefacts were reported 0 to 11 cases for CS-VIBE and 0 to 10 cases for TWIST. In the CS-VIBE cases, pulsation artefacts were seen only in the axilla. In 1 reader, these artefacts were significantly higher in CS-VIBE images (R3, P = 0.004, see Figure 9.2). Ghosting and infolding artefacts were reported in 0 to 6 cases for CS-VIBE and 0 to 13 cases in TWIST. In 1 reader, TWIST (R5, see Figure 9.3) significantly showed more of these artefacts (P = 0.039). One reader reported severe susceptibility artefacts in 4 cases for CS-VIBE and 6 for TWIST.

Table 9.2: Image quality reader study

	CS-VIBE					Inter-reader agreement[+]	TWIST					Inter-reader agreement[+]	Impact of artefacts (P-value per reader)[$]
	R1	R2	R3	R4	R5		R1	R2	R3	R4	R5		
Overall image quality score (mean)	75	66	70	69	56	0.112 (P=0.337)	70	63	76	71	45	0.656 (P<0.001)	N/A
Motion artefacts*						0.604 (P<0.001)						0.144 (P=0.010)	
1	0	0	0	0	1		0	1	1	0	3		R1: N/A
2	0	0	0	0	2		0	1	5	5	6		R2: 0.500
3	4	0	4	9	6		4	0	4	9	5		R3: 0.031
4	11	0	12	14	8		14	11	5	9	8		R4: 0.063
5	8	23	7	0	6		5	10	8	0	1		R5: 0.109
Breathing and pulsation artefacts*						0.155 (P=0.007)						0.211 (P<0.001)	
1	0	0	3	0	3		0	1	0	0	2		R1: N/A
2	0	1	6	6	8		0	2	0	1	8		R2: 0.625
3	15	3	11	14	10		3	6	4	12	6		R3: 0.004
4	8	18	3	3	1		13	14	13	10	6		R4: 0.063
5	0	1	0	0	1		7	0	6	0	1		R5: 0.999
Infolding and ghosting artefacts*						0.301 (P<0.001)						0.110 (P=0.052)	
1	0	0	0	0	3		0	0	1	0	7		R1: 0.500
2	0	2	0	4	3		2	4	1	2	6		R2: 0.999
3	1	12	1	15	6		11	14	3	8	6		R3: 0.500
4	10	8	1	4	8		9	4	3	13	2		R4: 0.625
5	12	0	21	0	3		1	1	15	0	2		R5: 0.039
Susceptibility artefacts*						0.528 (P<0.001)						0.412 (P=0.052)	
1	0	0	0	0	1		0	0	0	0	0		R1: N/A
2	0	0	0	0	3		0	0	0	0	6		R2: N/A
3	2	2	0	0	12		2	0	0	1	10		R3: N/A
4	21	2	2	21	7		19	3	1	14	5		R4: N/A
5	0	19	21	2	0		2	20	22	8	2		R5: 0.625

* Rated as 1: heavily disturbing to 5: not present; [+] Inter-reader agreement using Fleiss kappa was calculated on the dichotomized artefact values, for overall image quality ICC was used; [$] The impact of the artefacts was assessed on the dichotomized artefacts values, McNemar's test was used to assess the differences between the sequences per reader; N/A: not applicable, when no severe artefacts (score of 1 or 2) were present, no McNemar's test was performed

Reader Study 2: Lesion Appearance on CS-VIBE Compared with VIBE-DIXON

For this analysis, 31 lesions were re-evaluated (see Figure 9.4). Although lesion conspicuity scores were slightly lower for CS-VIBE (average CS-VIBE = 3.3, average VIBE-DIXON = 3.6, P = 0.004, see Table 9.3), sensitivity and specificity did not significantly differ (P = 0.500 and P = 0.327 for sensitivity and specificity, respectively). Pooled sensitivity and specificity were 84.0% and 60.0% for CS-VIBE and 92.0% and 55.4% for VIBE-DIXON. Intrareader agreement for BI-RADS descriptors was slight to almost perfect depending on the descriptor tested and ranged from 0.145 to 0.937 between sequences, as shown in Table 9.4. Overall intrareader variability was structurally less than interreader variability, hence showing more effect from readers than from the sequences used.

Table 9.3: Lesion assessment of lesion appearance reader study

	CS-VIBE					VIBE-DIXON				
	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5
Assessment category										
- BI-RADS 1	1	10	1	0	4	0	1	0	0	1
- BI-RADS 2	20	11	10	14	11	20	11	5	23	13
- BI-RADS 3	4	8	5	11	2	5	14	8	6	3
- BI-RADS 4	2	2	11	6	14	2	2	14	1	12
- BI-RADS 5	4	0	4	0	0	4	3	4	1	2
Lesion conspicuity score *										
1	0	12	1	0	4	0	7	0	0	1
2	3	10	6	0	2	2	7	2	1	2
3	7	9	11	14	8	12	10	8	9	6
4	17	0	9	7	7	14	7	13	12	5
5	4	0	4	10	10	3	0	8	9	17

*Lesion conspicuity is scored from 1: not visible to 5: very clear

Table 9.4: BI-RADS characteristics of lesion appearance reader study

	R1 CS-VIBE	R1 VIBE-DIXON	R1 Intra-reader agreement§	R2 CS-VIBE	R2 VIBE-DIXON	R2 Intra-reader agreement§	R3 CS-VIBE	R3 VIBE-DIXON	R3 Intra-reader agreement§	R4 CS-VIBE	R4 VIBE-DIXON	R4 Intra-reader agreement§	R5 CS-VIBE	R5 VIBE-DIXON	R5 Intra-reader agreement§	Inter-reader agreement+ CS-VIBE	Inter-reader agreement+ VIBE-DIXON
Lesions			0.936 (P<0.001)			0.145 (P=0.251)			0.613 (P<0.001)			0.526 (P<0.001)			0.424 (P=0.002)	0.295 (P=0.001)	0.331 (P<0.001)
- Focus	6	7		18	9		7	5		5	8		10	8			
- Mass	22	21		7	13		16	20		17	19		11	14			
- Non-mass	3	3		6	9		8	6		9	4		10	9			
Mass																	
Shape			0.867 (P<0.001)			0.571 (P=0.037)			1.000 (P<0.001)			0.575 (P=0.001)			0.323 (P=0.270)	N/A	0.350 (P<0.001)
- Round	2	0		2	5		2	3		3	4		4	3			
- Oval	17	18		4	5		10	13		9	10		5	9			
- Irregular	4	4		2	3		4	4		9	7		2	2			
Margin			1.000 (P<0.001)			0.308 (P=0.121)			0.605 (P=0.005)			0.590 (P=0.003)			0.235 (P=0.495)	N/A	0.383 (P<0.001)
- Circumscribed	19	18		6	10		11	13		13	15		8	8			
- Irregular	4	4		1	0		3	5		3	2		2	4			
- Spiculated	0	0		1	3		2	2		2	2		1	2			
Enhancement			0.861 (P<0.001)			0.444 (P=0.025)			0.450 (P=0.010)			0.206 (P=0.201)			0.462 (P=0.079)	N/A	0.322 (P<0.001)
- Homogeneous	10	8		4	9		12	13		15	17		4	6			
- Heterogeneous	11	12		3	2		3	4		2	2		3	5			
- Dark internal septation	1	1		0	1		0	2		0	0		0	1			
- Rim enhancement	1	1		0	1		1	1		1	0		4	2			
Non-mass																	
Distribution modifiers			0.400 (P=0.386)			1.000 (P=0.157)			0.400 (P=0.180)			1.000 (P=0.018)			0.750 (P=0.070)	N/A	N/A
- Focal area	2	1		3	4		3	2		4	2		5	5			
- Regional	0	0		1	1		0	0		3	1		1	0			
- Linear	0	0		1	1		4	2		0	0		4	4			
- Multiple regions	0	0		0	0		0	0		1	0		0	0			
- Segmental	0	0		1	3		1	2		1	1		0	0			
- Diffuse	1	2		1	0		0	0		0	0		0	0			
Internal enhancement			N/A			1.000 (P=0.157)			0.357 (P=0.266)			0.200 (P=0.505)			0.500 (P=0.248)	N/A	N/A
- Homogeneous	0	0		3	4		1	1		2	0		2	6			
- Heterogeneous	3	3		1	0		3	1		5	3		8	3			
- Clumped	0	0		3	5		4	4		2	1		0	0			
- Clustered ring	0	0		0	0		0	0		0	0		0	0			

§ Intra-reader agreement was assessed using weighted kappa statistics
+ Inter-reader agreement was assessed using Fleiss kappa, when there were too few complete cases for analysis, Fleiss kappa cannot be computed (N/A)
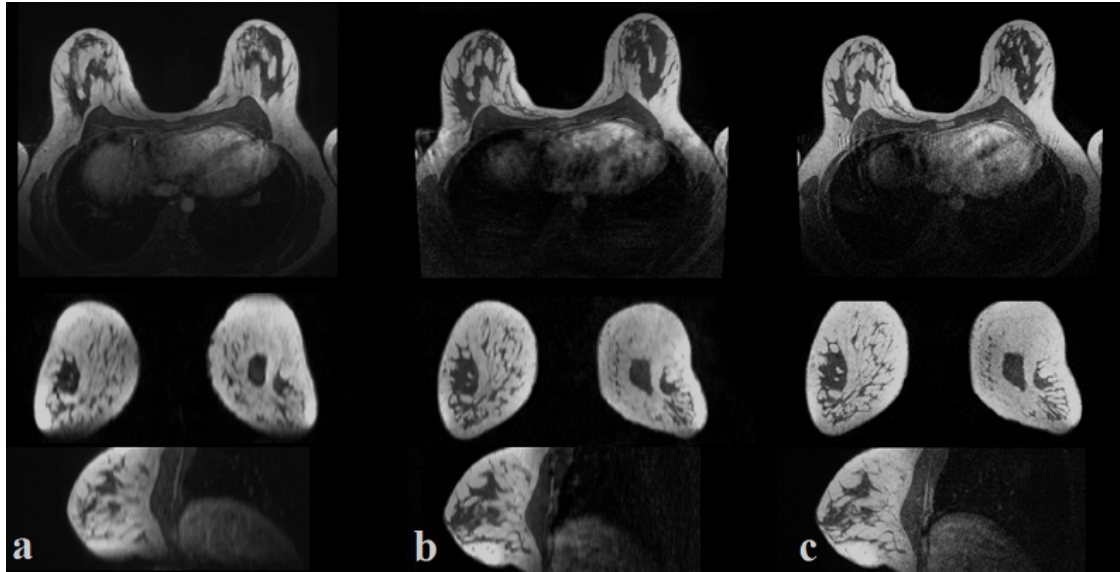
Figure 9.1: An example of the described sequences in a 28 years old woman. An example of TWIST of approximately one year earlier (a), CS-VIBE (b), and VIBE-DIXON (c) in axial (top), coronal (middle), and sagittal reconstruction (bottom).

Dynamic Analysis

Dynamic analysis (TTE and inflow curve type) was compared using 24 benign lesions in 11 women for which both CS-VIBE and TWIST acquisitions were available. The mean TTE was 16 seconds for the CS-VIBE and 22 seconds for TWIST (P = 0.448, sdCS-VIBE = 12.7 seconds, and sdTWIST = 9.2 seconds). Also, curve types were comparable between the CS-VIBE sequence and TWIST (P = 0.242). Most lesions showed a relatively slow wash-in of contrast (curve type 1), as is expected in benign lesions.

9.4 Discussion

In this study, we show that the use of compressed sensing allows strong improvement of the spatial resolution of ultrafast breast MRI. A reduction of more than 50% of the voxel size compared with a view-sharing approach is possible at comparable in-plane image quality. This allows multiplanar reformatting of axial acquisitions and thus morphological assessment of eventual lesions in 3 perpendicular planes. In fact, we found that morphologic assessment of specified lesions was not significantly different between CS-VIBE and the 18 times lengthier T1-weighted 3D Dixon sequence, and we also did not observe significant differences in sensitivity or specificity between these sequences, thus showing the potential of using ultrafast acquisitions alone for both morphologic and dynamic analysis.
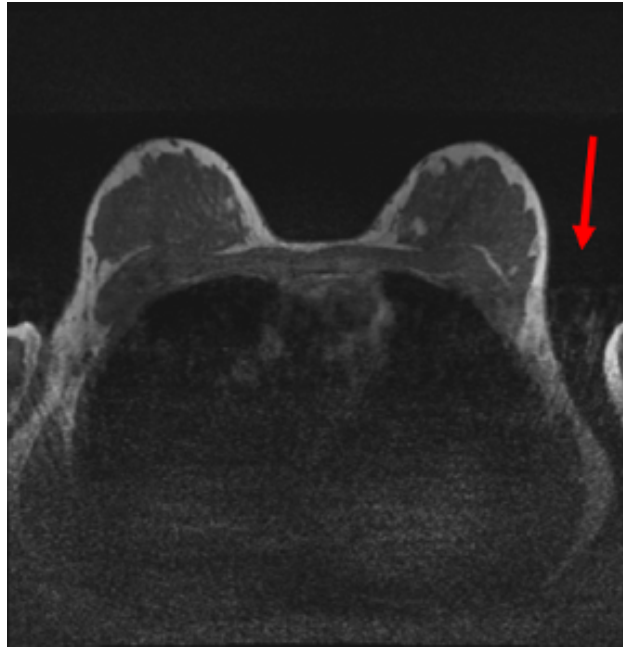
Figure 9.2: Example of a typical pulsation artefact in CS-VIBE; best seen in the left axillary region.
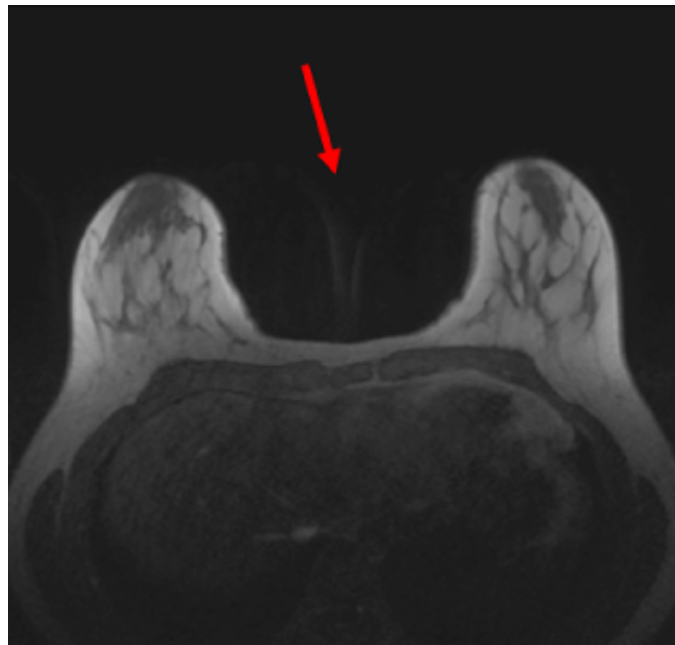

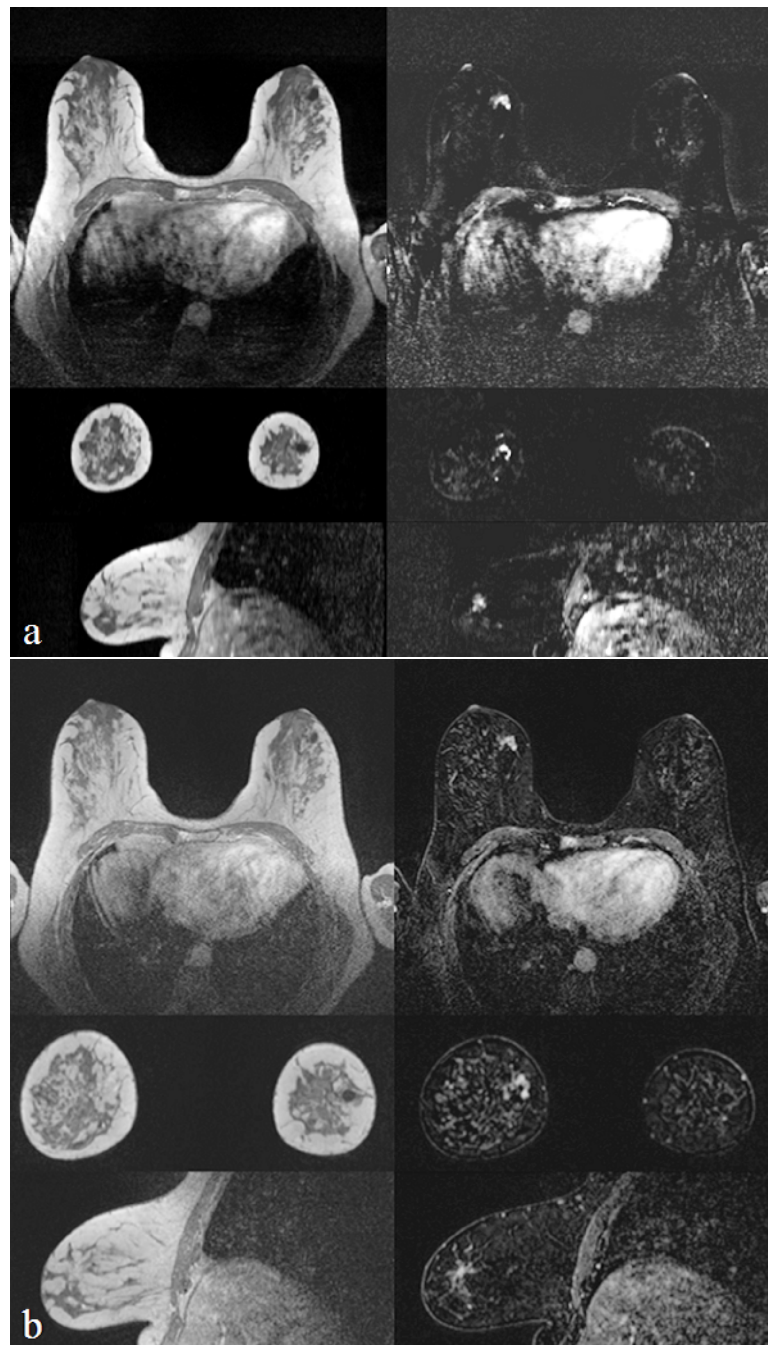**Figure 9.3:** Example of a typical ghosting artefact in TWIST

Figure 9.4: Lesion morphology in CS-VIBE (a, latest time-point (20)) and VIBE-DIXON (b, first time point) in a 53-year-old woman with an invasive ductal carcinoma in the right breast is virtually identical in all planes. Left column: original images; right column subtraction images.

Compressed-sensing volume-interpolated breath-hold examination seems to reduce ghosting and infolding artefacts compared with TWIST, although the results of the current study are not significant. This is important as Le et al[229] also noted that these artefacts were common in TWIST, and may impair image quality in the breasts.

In CS-VIBE, a specific pulsation artefact was commonly seen in the axillary region, where propagation of heartbeats sometimes distorts the image. Whether this is important when screening for breast cancer is an area of debate. Less than 1% of all breast cancers are detected by axillary metastasis alone[246], and the specificity of breast MRI for the detection of abnormal lymph nodes is only moderate at best[247]. Even in ultrasound screening, it was recently shown that assessment of the axilla mainly leads to more false-positive findings and is thus discouraged[248]. Consequently, in a screening setting, the assessment of the axilla seems to be of limited relevance, and artefacts in this region are therefore relatively acceptable.

Initially, ultrafast DCE-MRI acquisitions of breast lesions were developed to capture the inflow of contrast in breast lesions. The information obtained could be quantified using pharmacokinetic models that yielded the contrast transfer coefficient, the rate constant, and the relative fraction of the extravascular extracellular space. It was shown in several studies that this information could be used to improve the discrimination between benign and malignant lesions[249]. In terms of spatial resolution, the sequences used for this purpose were poor, and therefore morphologic evaluation of lesions using these images alone was long deemed impossible. Only with the development of faster view-sharing sequences such as TWIST, 4d TRAK, and DISCO, it became possible to obtain inflow information at a spatial resolution that was considered diagnostically adequate[55,230]. Simultaneously, alternative approaches for interpretation of the dynamic data, such as maximum slope and time to enhancement, that abandoned the quantitative calculations, improved the usability of inflow dynamics for clinical practice. It was, thus, postulated that using TWIST alone, breast screening with MRI could be performed in less than 2 minutes[55]. Nonetheless, protocols now in use and under development for abbreviated breast MRI still use a hybrid approach[9], where the ultrafast series is interleaved between 2 high-resolution acquisitions during inflow of contrast. In practice, this adds about a minute of scan time to the dynamic protocol, which may not seem much but is an increase in overall scan time of approximately 150% and likely also adds to the interpretation time. Further improvement of ultrafast acquisitions that completely resolve the need for additional high-resolution acquisitions is thus warranted, and CS-VIBE seems an important step in this process.

In addition to the limited through-plane resolution of TWIST, all view-sharing sequences also suffer from temporal blurring. This may cause a delay in the visualization of small vessels and reduce the discriminative power of enhancement charac-

teristics. Rapacchi et al[250] combined compressed sensing with TWIST to reduce the temporal footprint. This reduced blurring due to motion in thoracic MRA and increased overall vessel sharpness. In our study, we showed that dynamic parameters obtained with TWIST and CS-VIBE were comparable for benign lesions. However, due to the low number of malignant cases in our dataset (N = 5), we were unable to determine the discriminating capacity between benign and malignant breast lesions of the dynamic parameters obtained from the CS-VIBE. The reduction in temporal blurring might imply that cutoffs determined for view-sharing approaches should be slightly adjusted for the compressed sensing approach. Future research should determine whether there is a significant difference in time to enhancement and maximum slope for TWIST and CS-VIBE and whether this may have a clinical impact on lesion management.

There are some limitations to our study. In the current study, apart from the presence of artefacts, we did not quantify other image quality aspects such as sharpness of contours or image homogeneity but rather focused on the subjective reader assessment of image quality for diagnosis. For quality comparison, we used TWIST images that were obtained approximately 1 year before the CS-VIBE acquisitions, rather than performing both sequences in succession in the same patient. This may have increased the presence of positioning differences and associated motion artefacts[251]. However, we thought this was appropriate due to the fact that the value of ultrafast acquisitions can only be appreciated during contrast administration. Performing the 2 acquisitions in succession would require a contrast wash-out period in between and thus patients would anyway be taken out of the scanner and repositioned for the next acquisition. Differences in positioning and movement artefacts would therefore likely have occurred in a similar fashion using this approach. This would, however, have prevented eventual differences in breast shape and volume caused by weight gain or loss of our patients. Nonetheless, giving contrast twice to the same patient only for comparison of 2 sequences might be regarded as less acceptable in light of the current discussion on the safety of gadolinium-based contrast agents[86]. This holds also true for the comparison of CS-VIBE versus VIBE-DIXON, in which, due to the structural differences in timing (the VIBE-DIXON was always acquired after the CS-VIBE) different levels of background parenchymal enhancement were observed. In theory, this may slightly affect the morphological description of the lesions.

Furthermore, we investigated multiple lesions within the same patient, thus assuming that all lesions can be considered as independent primary lesions. Although this

assumption likely holds true for benign lesions, it may not be completely valid in malignant lesions as multicentric or multifocal disease is present in 20% to 25% of the malignant cases[11]. To overcome this issue, we only evaluated lesions that were reported as independent abnormalities in clinical evaluation. This implied that, only in 1 patient with a malignancy, 2 malignant lesions were scored, as these were highly different in their morphological appearance.

As commonly reported[252], interreader variability of morphological BI-RADS descriptors was large. In this specific study, it might have even been larger than usually reported due to the different clinical background and experience of our readers. However, because we mainly focused on the effects of the different sequences on intrareader variability, this is of less concern.

A current limitation for implementation of the CS-VIBE is that, on the standard scanner hardware, reconstruction of the complete time series per subject still takes 45 minutes (depending on breast size). Currently, this is impairing clinical workflow. However, future optimization of the reconstruction process will probably render the technique clinically viable. In addition, it is possible to perform the reconstructions offline, which now already allows a clinically viable pipeline, considering that it seems acceptable that scans are read a day after acquisition.

In conclusion, the IQ of CS-VIBE was comparable to TWIST while a much higher spatial resolution can be achieved. In addition, CS-VIBE reduces temporal blurring in the evaluation of inflow dynamics and enables multiplanar assessment of lesion morphology. Compressed-sensing volume-interpolated breath-hold examination is a promising technique for ultrafast breast MRI with high spatiotemporal resolution.

General discussion

10

The early detection of breast cancer has proven to increase survival[253–255]. Mainly because of this finding, many countries have introduced population-based mammographic screening programs. However, mammography has its limitations. These limitations make mammography less suitable as a stand-alone technique for screening of women at high risk of developing breast cancer. Women at increased risk for breast cancer are, therefore, invited for intensified screening programs composed of a yearly breast MRI in addition to a yearly mammography exam.

Prospective trials[14,15,37–42] have evaluated the performance of these breast MRI screening programs. In these trials, the screening performance was measured by several performance measures that are regularly used for auditing purposes. In accordance with the terminology used for mammographic screening: recall rate, biopsy rate, positive predictive value of recall, positive predictive value of biopsy, cancer detection rate, sensitivity, and specificity are the usually reported metrics. Sensitivity and specificity numbers of breast MRI reported by prospective trials are in the range of 71% to 91% and 89% to 97%, respectively[14,15,37–42].

The detection and appearance of most breast cancers on breast MRI are relatively straightforward. Cancers can be recognized by their fast and strong enhancement as they typically light up within a minute after the contrast agent is injected. However, not only cancers light up with breast MRI, benign lesions or even normal breast parenchyma can also have a relatively fast enhancement with wash-out. This is the reason why breast MRI has a lower specificity compared to mammography[15].

Before the publication of several articles in this thesis, only data on prospective clinical trials in sharply defined population-groups were reported. Data on the real-world performance of breast MRI screening programs for women at increased risk in clinical practice was lacking.

10.1 Screening performance of breast MRI and mammography

In the first chapters of this thesis (Chapter 3 - Chapter 5), the current breast MRI screening program at our institution was analyzed. It was shown that the overall program has a relatively good performance. Nevertheless, in women with certain risk factors, the sensitivity is still far from perfect. The data, presented in Chapter 3, show that in *BRCA1* mutation carriers the sensitivity of our breast MRI screening

program was still inadequate and was as low as 81.3% (95% CI: 68.1 - 89.8%). This means that in these women, both mammography and breast MRI have a suboptimal performance. Furthermore, sensitivity is excellent (95.2%, 95% CI: 74.1 - 99.8%) in women with a family history of breast cancer without a proven genetic predisposition. However, the frequency of biopsy for benign findings is relatively high (PPV3 of 0.14 (95% CI: 0.09 - 0.21)). Overall, Chapter 3 clearly shows that breast MRI has a superior performance compared to mammography for screening these high-risk women. As previously mentioned, according to skeptics, one of the disadvantages of breast MRI is the low specificity. However, our results showed that breast MRI appears to be only slightly lower than that of mammography in the follow-up rounds.

Resulting from the statements above, the question whether mammography is of added value when breast MRI is available follows logically. Certainly, some cancers will be detected solely by mammography. However, these cancers are mostly found to be of low grade or even in situ cancers only, which can be characterized as cancers with a relatively good prognosis. Some argue that mammography may cause overdiagnosis in population-based mammography screening. Estimates of overdiagnosis range from 1% to 30% of screen-detected cancers[256–259]. In fact, mammography has a technology-inherent bias to detect preferentially slowly growing cancers[75]. Mammography highlights particularly architectural distortions and calcifications, which are generally signs of slow-growing breast cancers of low grade[260]. The rate of overdiagnosis of MRI-only detected cancers is currently unknown. However, based on the grounds for enhancement and thus detection, MRI is biased in the opposite way and preferentially detects biologically active cancers. This is in accordance with studies that show that especially high-grade DCIS is better detected by breast MRI than by mammography[18]. Consequently, MRI might be less susceptible to overdiagnosis than mammography.

In Chapter 4, the limited value of mammography when breast MRI is available is discussed in more detail. Only 13 cancers were detected by mammography only, of which most were detected in women of age 50 or older (77%) and only 3 were found in *BRCA* mutation carriers. In addition, most of these cancers were in situ cancers only (62%). Moreover, mammography showed to increase false positives. Results showed that, while in older women the additional cancer detection increases, there is also an increase in false positive findings. Although the number of cancers detected in our hospital is low, our results strongly suggest that mammography is not of added value before the age of 40. We might even go further and increase this age with another five years (to 45 years) according to the results of our dataset for

women carrying a *BRCA* mutation. However, to give a substantiated recommendation, it would be most optimal to combine data from several hospitals and create a large dataset to be able to identify a well-defined nationwide (or even worldwide) consensus on the age to start with mammography when breast MRI is available.

To conclude, in clinical practice, the high-risk screening program is performed in a 'one-size-fits-all' strategy. Based on our results, this is not satisfactory for all women, as risks, cancer type and aggressiveness may vary, and imaging modalities such as mammography and MRI have different sensitivities between subgroups. Personalized screening programs are therefore increasingly investigated and will likely represent the future of breast cancer screening programs. This personalized screening might result in more intensive surveillance for some women, for example by performing biannual screening tests in *BRCA1* mutation carriers, hopefully resulting in higher cancer detection rates, while, simultaneously, personalization of screening practice might decrease the burden of intensified screening in other women.

10.2 Breast cancers on prior MRI scans

In Chapter 5, we present and discuss the results of our study in which we investigated the visibility of breast cancers detected in breast MRI screening on prior MRI-scans. This showed that almost one-third of the breast cancers that were detected could have been detected at an earlier point in time (6 - 24 months prior to detection). These cancers were missed or misinterpreted in clinical practice. From studies on mammography screening, it has become clear that in mammography similar numbers are missed in screening, and that this is one of the reasons why second reading is introduced for mammography screening. In fact, one-third of the cancers detected in mammography screening could have been identified at an earlier scan when retrospectively evaluating the same scans[261]. These results can be explained by the fact that human beings, who are subject to error, perform these tasks. However, one has to also consider the possible bias when re-evaluating the scans with knowledge of outcome, which was the case in our study. A definitive conclusion of why radiologists missed or misinterpreted breast MRI-visible cancers is currently unknown. Thus, to be able to distinguish the difference between missed and misinterpreted lesions, additional research using eye tracking is needed. In other research areas, this is already evaluated. An observer study employing eye tracking found that the majority of errors in a cancer detection task in chest radiographs were due to misinterpretation and that very few nodules were actually missed by search errors[262]. However, such studies are usually conducted as reader studies with highly

enriched datasets and may not reflect true clinical practice in which the amount of possible disturbances is much higher. Consequently, future work needs to analyze whether lesions are missed in breast MRI due to oversight or misinterpretation errors.

In addition, it is essential to assess whether there are specific morphological or other features that predispose for misinterpretation errors. When this is the case, training of breast radiologists on these specific features might already largely reduce the frequency of such mistakes. Furthermore, it appears interesting to assess the consequences of the current errors; i.e. what is the growth of the missed cancers, does this imply a stage shift? Is there a difference in the frequency of axillary spread between cancers visible in prior scans and those that are truly new? And so on. This remains future work for the moment. In any case, to improve early breast cancer detection with breast MRI, it appears essential to reduce reading errors. Based on our study, we propose to introduce second reading for breast MRI, as is commonly employed for screening mammography. However, the actual gain of this procedure in terms of sensitivity still needs to be assessed. This requires a new study in which the breast cancer cases are mixed with a large number of normal scans in order to document the magnitude of the benefit, which was beyond the scope of the studies we performed here. An alternative to second reading would be the implementation of a computer-aided detection (CAD) system for breast MRI. A prototype of such a CAD system for breast MRI was developed by Gubern-Merida et al.[132]. This CAD system was applied on a subset of missed cancers from our study. The results showed that the CAD system could detect 63% of the initially misinterpreted/ overlooked cases at a rate of 4 false positive annotations per normal scan. Applying this CAD system as a first reader, showing the most suspicious areas, might, therefore, help in detecting breast cancers in breast MRI and possibly making increasing the sensitivity and cost-effectiveness. However, further improvements are still necessary before this can be used in clinical practice as the very high false positive rate will likely strongly curb the enthusiasm of reporting radiologists.

10.3   BPE and breast cancer

Since the introduction of the new BI-RADS MR lexicon in 2013, background parenchymal enhancement (BPE) is a characteristic that radiologists need to score on breast MR images. BPE describes the amount of enhancement of the normal fibroglandular tissue at approximately 90 seconds after contrast injection (i.e. at the time of peak enhancement in most lesions). In the current version of BI-RADS, BPE is divided into 4

categories which are visually assessed: minimal, mild, moderate, and marked. From previous research, it is already known that BPE is correlated to age and menstrual cycle[158,200,201]. Since the introduction of breast DCE-MRI BPE is a topic of debate. Some argue that higher BPE scores are correlated to abnormal BI-RADS scores (BI-RADS 3 or higher) or even increase false positive biopsies[159,174,213]. Recent studies have shown that higher BPE might be correlated to increased breast cancer risk[23,24], thus showing a potential to include BPE as a risk factor to personalize screening programs.

In order to use BPE as a stratification parameter, it should, however, be independent of the type of cancer in the breast, or should preferentially select the more aggressive cancers. However, our findings in Chapter 6 show that BPE might be negatively correlated with tumor grade and progesterone negativity. Thus, introducing BPE as a risk factor for breast cancer, as described by others[23,24], could potentially lead to the detection of the more indolent cancers and leave the fast-growing cancers undetected. This points to a different pathogenesis in breast cancer development where hormonal stimulation might be of less importance in aggressive cancers than in relatively indolent cancers. However, since exact causes for fluctuations in BPE are largely unknown, it should be extensively investigated before using it to stratify screening protocols.

Moreover, as previously mentioned, recent publications showed that BPE correlated to breast cancer risk in a case-control study design. These results were completely or partly based on breast MRI examinations of breast cancer patients. This means that these studies used BPE more as a diagnostic tool, rather than a predictive one. To be able to implement BPE in risk prediction tools, we need to unravel the predictive value of BPE. In Chapter 7, we tried to predict breast cancer in our high-risk population using already known independent risk factors (age and *BRCA* status) and added amount of fibroglandular tissue (FGT) and BPE measures (calculated on breast MR images) to this model. For this model, the first breast MRI (baseline breast MRI) was used to measure FGT and BPE values. To account for the large inter-reader variability reported previously for BPE, in this particular study, we used an automated deep learning method, presenting discrete values for FGT and BPE and created well-defined cut-offs between classes. This method was based on methods described in Dalmis et al.[105]. We discovered that, when investigating baseline MRI scans, FGT and BPE had no predictive value for breast cancer risk and therefore cannot be used to stratify a screening program for high-risk patients. In fact, based on our findings in Chapter 6, any stratification based on these factors would lead to a less favorable

outcome. In line with some earlier reports, our results showed that higher amounts of FGT and BPE increase the risk of false positives in the first round of screening. Fortunately, neither FGT nor BPE is predictive of false positive findings in future screening rounds.

Although our results are important for clinical practice, this study was based on retrospective data. Thus, further validation is needed by performing a prospective multi-center trial, using matched hardware and protocols.

10.4   Compressed Sensing for breast MRI

Currently, breast MRI is only proven to be cost-effective in the highest breast cancer risk categories[165,263]. However, in clinical practice, and partly based upon guidelines, it is being applied to many women in lower risk categories, as already described in Chapter 3 of this thesis. To increase the cost-effectiveness of breast MRI and to be able to offer it to a larger population, it would be beneficial to shorten breast MRI protocols. Breast MRI protocols can now take approximately 20 minutes to perform. Several research groups showed that shortening the protocol, creating an abbreviated protocol (by excluding sequences from the protocol), might still reach the same cancer detection rates, albeit at a cost of a slight increase in false positive findings because valuable classification information is discarded[16,17,53,76,223–225,237,238]. Although first results are promising, further research in real screening populations is needed to obtain more definitive results. An overview of the current abbreviated protocols used is provided in Chapter 8.

Mann et al.[55] showed that wash-in information, acquired with an ultrafast sequence (TWIST), had a better discriminative power (benign versus malignant) than wash-out information acquired with the standard T1 sequence, and might, therefore, compensate the lack of classifying information in abbreviated protocols. However, the downside of the ultrafast sequence used by Mann et al.[55] is that it relies on view-sharing (data-points are re-used in several time points) and has a low through-plane resolution. Hence, we tested a new sequence (using a compressed sensing approach), which has the advantage of being ultrafast at a high in-plane and reasonable through-plane (0.8 x 0.8 x 1.6 mm) spatial resolution, without the view-sharing disadvantage. In Chapter 9, it was shown that this new sequence provided morphological information that was comparable to VIBE-DIXON: our regular T1-weighted 3D-DIXON sequence; and provided similar dynamic information as TWIST: the ultrafast view-sharing sequence currently in use in our hospital. This shows that it

might be possible to shorten breast MRI protocols by a factor five, without losing any (or almost no) information. This would make breast MRI less expensive and thus more accessible to the general public. However, one has to keep in mind that this study presented in Chapter 9 was a proof-of-concept study including only 30 women. A larger trial is needed to provide more information on the applicability in clinical practice.

10.5   Consequences of this thesis for daily practice

This thesis provides us with valuable knowledge on the performance of a high-risk screening program in clinical practice. Although the population investigated in this thesis was one of the largest ones in literature, even larger datasets are desirable to better document realistic confidence intervals around the given performance measures. Nonetheless, our results show that the sensitivity of the entire screening program is still poor in *BRCA1* mutation carriers, which might lead to additional screening for these women. Furthermore, the screening program showed significantly lower specificity values for women with only a family history of breast cancer. This indicates that strict evaluation of indication and tailored patient information is required for intensified screening. A meta-analysis of similar studies (as is available for *BRCA1* mutation carriers[96,111]) is needed to provide us with figures that can be used as an initial benchmark for auditing of MRI screening practice. However, even in the absence of such figures, regular self-reported quality control of breast MRI screening outcomes might be advisable.

Concerning the younger women in our increased risk population, current guidelines recommend that all women from the age of 30 need to undergo a yearly mammography. However, as our results showed in Chapter 4, we can conclude that this might be too early and these mammographies do not add in cancer detection. These only add in false positives and thus costs, but also leading to unnecessary anxiety for the screened women. Therefore, we propose to revise these guidelines. Fortunately, new national guidelines are already pending and will raise the age at start of mammography for *BRCA1* mutation carriers.

Despite the important task of detecting breast cancers in women with a high chance of getting breast cancer with a bad prognosis, radiologists tend to miss or misinterpret one out of three breast cancers. While this may be a natural consequence of the wish to optimize both sensitivity and specificity, we can conclude that there is still much to gain. Future implementations of second reading, or, likely better, computer-

aided diagnosis (CAD) could improve detection rates and/or reduce false positives. Nonetheless, future research is needed to increase the specificity of these CAD systems. Subsequently, we must also ensure that CAD systems are implemented in the most optimal way, as multiple approaches for the implementation of CAD systems into clinical practice are possible (e.g. image enhancement, first reader, second reader, arbiter), as briefly described by Gubern-Merida and coworkers[132]. Depending on the way CAD is implemented the added value of CAD will differ.

As previously mentioned, there is a general shift towards personalized medicine, which is also seen in screening regimens using personalized risk models to stratify women to different screening protocols. In the general population, research is ongoing for women in the densest breast category (extremely dense breasts), introducing breast MRI as a supplemental screening technique as mammographic sensitivity is known to be lacking. However, for high-risk screening, our results show that BPE and amount of FGT cannot be considered as independent risk predictors for breast cancer, and should therefore not be used for further stratification of women who are already at an increased risk for breast cancer based upon other factors. Our study on ultrafast breast MRI using compressed sensing provides further evidence that the technique is nearly ready for large-scale implementation in breast MRI screening practice and might largely reduce the costs of these programs, making implementation in other risk groups, such as the above-mentioned group of women with extremely dense breasts, more viable.

10.6 Future perspectives of breast MRI screening

Despite the fact that this thesis partly focuses on increasing the cost-effectiveness of breast MRI by implementing shorter protocols, cost-effectiveness is not the only drawback of breast MRI. All breast MRI protocols rely on intravenous administration of a gadolinium-containing contrast agent. Until recently, gadolinium-based contrast agents were considered to be amongst the safest contrast agents in use. However, due to recent findings of gadolinium deposits in the brain, these contrast agents are now a much-debated subject, as already briefly mentioned in Chapter 2. Since there are no documented clinical consequences of these depositions, both the International Society of Magnetic Resonance in Medicine (ISMRM) and the European Society of Breast Imaging (EUSOBI) still endorse all current indications for breast MRI. However, the discovery has fueled the debate on ethical considerations for breast MRI as a screening modality, where healthy women will be repetitively exposed. Since it might take extremely long to uncover eventual effects of gadolinium

deposition, this debate will not be resolved on short notice. As either abbreviated, ultrafast, or hybrid protocols still require contrast, these techniques do not solve this problem. This is an incentive to increase the efforts in investigation of alternative breast MRI strategies that do not rely on intravenous contrast administration, such as diffusion-weighted techniques (including diffusion tensor imaging (DTI) and intravoxel incoherent motion (IVIM)) and non-contrast vessel analysis (e.g. arterial spin labeling (ASL)) or using other imaging modalities that do not rely on contrast administration. Recent technological developments in cancer imaging have already led to a shift towards functional assessment of tissue characteristics, hopefully making contrast administration in the (near) future unnecessary. Further development of CAD for detection and classification of lesions in non-enhanced breast MRI might allow for capabilities beyond those of human perception and therefore allow the use of sequences where we as humans are not good enough to determine the presence of cancer.

10.7 Conclusion

In conclusion, this thesis adds essential knowledge on the effectiveness of the (MRI) screening program for women at increased risk for breast cancer and opens several roads to improvement of current practice.

Summary

11

Breast cancer is one of the leading causes of cancer death among women. In the Netherlands, one out of 6.6 women faces the diagnosis of breast cancer during her life[1]. The survival chances are increased by the early detection of breast cancer[1,2]. To ensure early detection, breast cancer screening was introduced in the Netherlands, in which women from the age of 50 are invited for a biennial mammography. Unfortunately, mammography has its limitations, causing that breast parenchyma can obscure cancer. In addition, research has already proven that high breast density is an independent risk factor for breast cancer[7]. Other significant risk factors are personal history of breast cancer, radiation to the chest at a young age, extensive family history of breast cancer, and a proven *BRCA* gene mutation[12].

Women with a lifetime risk of $\geq$ 20 - 25% for breast cancer are, depending on local policies, invited for intensified screening programs. In these screening programs, women are yearly screened with breast MRI and mammography. In the case of a known *BRCA* mutation, screening starts with MRI from the age of 25 and from the age of 30 a mammography is added to this regimen.

Breast MRI has been proven to be the most sensitive imaging modality for the detection of breast cancer. However, the actual performance of breast MRI for screening of women at increased risk for breast cancer in clinical practice is unknown. This thesis focuses on breast MRI as a screening tool and the performance of this intensified screening program. Furthermore, we assess how the current regimen can be improved. Commonly used protocols for screening breast MRI are described in detail in Chapter 2.

Chapter 3 describes the performance of the screening program for women with an increased risk for breast cancer in detail. In this Chapter, different risk categories were taken into account. The performance measures that were used are based on those usually used in audits for mammography screening[33]. This study shows that, despite the fact that breast MRI is the most sensitive imaging modality for breast cancer detection, breast MRI only had a moderate performance in women with a *BRCA1* mutation (at a sensitivity of 81.3% for the complete screening program). Additionally, in women with only a family history of breast cancer, the breast cancer detection rate was far below the average of the other risk categories (7.6 per 1000 exams versus 14.6 per 1000 exams), whereas the positive predictive value of biopsy was substantially lower. From this data, we can conclude that more personalized screening programs fit better for these risk categories.

In Chapter 4, the added value of mammography when breast MRI is available was investigated in more detail. From previous research is known that, for women with a high breast density, mammography has a low sensitivity[122]. Besides that, women participating in high-risk screening programs are relatively young (which is one of the reasons for having high-density breasts). In this Chapter, the influence of mammography on false positives (recalls and biopsies) and cancer detection were investigated in two risk categories (*BRCA* mutation carriers and others) and several age categories. We concluded that mammography had only limited value for cancer detection in this high-risk screening population. In total, 13 of the 125 cancers were solely detected with mammography. The results show that mammography was only of value in older women; in the group of women below the age of 40, mammography was of no added value for cancer detection. Of the 13 women with solely mammography detected breast cancer, most women were approximately 50 years of age and many of them had only in situ disease (DCIS, 8/13). Thus, we want to emphasize that the age for the start of mammography could be increased to at least 40 years, as opposed to the current practice of starting at the age of 30. This would also spare women the negative side effects of these mammographic studies such as increased radiation exposure.

Since the primary goal of screening is early detection, it is of great importance that cancers are detected at the proper time. That is the reason why Chapter 5 was written. In this Chapter, we investigated whether breast cancer detected in the intensified screening program (screen-detected cancers, interval cancers, incidental cancers detected in mastectomy specimen) could have been detected at an earlier screening round. In the assessment of the prior MRI scans, we made the distinction between visible, minimally visible and invisible cancers. Our results showed that almost one-third (31%) could be considered as missed cancers since these were considered as visible and suspicious on the prior exam by re-evaluation. In addition, 34% of the breast cancers were scored as minimally visible on the prior MRI scan. The reason why these breast cancers were not detected in clinical practice can be attributed to two factors: 1. simply not noticing the breast cancer, or 2. the misinterpretation of lesions. This needs further investigation in future research.

High-risk screening programs are currently performed using a 'one-size-fits-all' strategy, in which everyone eligible to intensified screening receives the same screening exams. However, not everyone has the same breast cancer risk and risks can change during a woman's life. This is one of the reasons why currently much research is focused on personalizing screening programs. In Chapter 6 and 7 we investigated

the possibilities for further personalization of screening strategies in women at increased risk. In detail, we investigated whether two potential features of the normal breast measured on breast MRI can be used as risk prediction tools. The fraction of fibroglandular tissue (FGT) within the breast is a well-known risk factor for women at average risk. The more fibroglandular tissue, the higher the risk. Background parenchymal enhancement (BPE) is the enhancement of the normal breast parenchyma. King et al.[23] speculated that this factor (BPE) could potentially help in the stratification of women to more or less intensified screening regimen as they observed higher rates of BPE in the contralateral breast of women with breast cancer than in women without breast cancer. This finding led to the performance of the study described in Chapter 6, where we re-evaluated the contra-lateral breast MRI of women with breast cancer and correlated these BPE scores to tumor characteristics. Our results showed that the pathological tumor grade was negatively associated with BPE scores, which means that in women with high background enhancement the chances were higher to detect a cancer of lower grade with a better prognosis than in women with lower scores of BPE. Thus, when we adhere to the recommendations of King et al.[23], this might lead to the detection of more indolent cancers and might leave the aggressive cancers undetected. With this Chapter, we show that great care must be taken before using potential predictors for patient stratification. In Chapter 7, the potential risk predictors that can be extracted from breast MRI data (FGT and BPE values) are investigated in more detail. In this Chapter, we created a risk prediction model for future breast cancer occurrence. We used a large dataset of 1533 breast MRI examinations. In order to obtain quantitative and reproducible BPE and FGT values, automated segmentation with a convolutional neural network (deep-learning) was used. BPE and FGT as measured on baseline MRI (at the start of screening) were used in the calculations. The results showed that neither BPE nor FGT values could predict future breast cancer occurrence in the high-risk population tested. However, both were related to the risk of false positives (both recall and biopsy) in this first round of screening.

One of the merits of personalization of a screening protocol is that it might increase the cost-effectiveness of screening programs. However, there are also other approaches to increase the cost-effectiveness of breast MRI screening. A very simple approach is to reduce the time it takes to obtain a single breast MRI examination. Such reduced MRI protocols are commonly referred to as 'abbreviated' breast MRI. The current evidence regarding such abbreviated protocols is extensively reviewed in Chapter 8. Most studies report a virtually equal sensitivity of abbreviated and normal scan protocols albeit sometimes at the cost of a small reduction in specificity.

Ultrafast breast MRI forms a special branch within the abbreviated protocols, as it still enables dynamic evaluation of contrast inflow within the lesion and, therefore, has a certain advantage over the other abbreviated protocols that rely solely on lesion morphology. Unfortunately, ultrafast breast MRI cannot be performed at equal spatial resolution as other abbreviated approaches and re-uses data-points for multiple time points, a process referred to as view-sharing, which leads to temporal blurring. In Chapter 9, we investigated the performance of a new sequence for breast MRI (CS-VIBE). This sequence that is also intended for ultrafast breast MRI, is based on a technique called 'compressed sensing', which can be described as a technique for efficiently obtaining and reconstructing the MR signal by finding a solution for underdefined linear systems[243]. Using this new sequence (CS-VIBE), we were capable of obtaining a full dynamic breast MRI scan in less than 2 minutes at a spatial resolution of 0.8 x 0.8 x 1.6 mm, without the use of view-sharing. In this 'proof-of-concept' study, we compared the image quality and lesion appearance of this new sequence to a view-sharing (TWIST) and regular T1-weighted examinations (VIBE-DIXON), showing that there was no significant difference between the new sequence and the TWIST regarding dynamic interpretation, whereas the spatial resolution of the new sequence was as good as that of the regular VIBE-DIXON sequence. CS-VIBE might, therefore, be a valuable alternative for the sequences currently in the protocol. This consequently brought us closer to "cheap" breast MRI examinations, enabling the use of breast MRI for screening in a larger group of women.

In summary, this thesis shows the current performance of an intensified screening program for women with an increased risk for breast cancer in an academic hospital. Several weaknesses have been identified and potential points for improvement have been discussed, for example by pointing out the lack of added value of routinely obtained mammography. Moreover, potential risk factors for personalized screening are investigated and criticized. Lastly, this thesis takes a quick glance at improving the cost-effectiveness of breast MRI for screening. In conclusion, this thesis offers handles to improve breast MRI screening programs. However, future research is needed to validate the results that have been mentioned in this thesis.

Samenvatting

12

Samenvatting

Borstkanker is de meest voorkomende doodsoorzaak door kanker bij vrouwen. Eén op de 6.6 vrouwen in Nederland krijgt te maken met borstkanker tijdens haar leven[1]. De overlevingskans wordt vergroot door het vroegtijdig opsporen[1,2]. Om dit mogelijk te maken is er in Nederland het mammografie screeningsprogramma ingevoerd, waarin vrouwen vanaf het 50e levensjaar tweejaarlijks worden uitgenodigd voor een mammografie onderzoek. Helaas kent mammografie zijn beperkingen, zo kan namelijk borstklierweefsel de tumor verbergen. Bovendien heeft wetenschappelijk onderzoek aangetoond dat borstklierweefsel met een hogere dichtheid (de aanwezigheid van veel borstklierweefsel) op zichzelf al een risicofactor is voor het ontwikkelen van borstkanker[7]. Andere significante risicofactoren voor het ontwikkelen van borstkanker zijn: persoonlijke geschiedenis van borstkanker, bestraling van de borst op jonge leeftijd, uitgebreide familiaire geschiedenis van borstkanker en de aanwezigheid van *BRCA* gen mutaties[12].

Afhankelijk van lokaal beleid worden vrouwen met levenslang risico van $\geq$ 20 - 25% om gedurende hun leven borstkanker te ontwikkelen uitgenodigd voor uitgebreidere screeningsprogramma's. In deze programma's worden de vrouwen uitgenodigd voor een jaarlijkse borst MRI en een mammografie. In het geval van een vrouw met een *BRCA* mutatie start screening met MRI vanaf het 25ste levensjaar en vanaf het 30ste levensjaar wordt mammografie aan dit regime toegevoegd.

Borst MRI wordt beschouwd als de meest sensitieve beeldvormende techniek voor het vroegtijdig detecteren van borstkanker. De daadwerkelijke effectiviteit in de praktijk van het borst MRI screeningsprogramma voor vrouwen met een verhoogd risico op borstkanker is niet bekend. Dit proefschrift focust op borst MRI als screeningsmiddel en de prestatie van het uitgebreide screeningsprogramma met borst MRI en mammografie. Verder wordt in deze thesis geëvalueerd hoe het huidige regime kan worden verbeterd. Veel gebruikte protocollen voor borst MRI worden in detail beschreven in hoofdstuk 2.

Hoofdstuk 3 beschrijft de prestatie van het uitgebreidere screeningsprogramma voor vrouwen met een verhoogd risico op borstkanker in detail. Hier wordt rekening gehouden met de verschillende risico categorieën. De uitkomstmaten die worden gepresenteerd zijn gebaseerd op de uitkomstmaten die gebruikt worden in audits van mammografie screeningsprogramma's[33]. Dit hoofdstuk laat zien dat, ondanks het feit dat borst MRI de meest sensitieve beeldvormende techniek is voor het detecteren van borstkanker, het maar matig presteert in vrouwen met een *BRCA1* mutatie (met een sensitiviteit van het complete screeningsprogramma van 81.3%). Verder

wordt ook duidelijk dat de borstkanker detectie in vrouwen met alleen een familie geschiedenis met borstkanker ver onder de andere risico categorieën ligt (7.6 per 1000 onderzoeken versus een gemiddelde van 14.6 per 1000 onderzoeken), terwijl de positief voorspellende waarde voor biopsie veel lager ligt. Uit deze data kan worden opgemaakt dat gepersonaliseerde screeningsprogramma's beter zouden passen in deze patiëntenpopulatie dan deze 'one-size-fits-all' strategie.

In hoofdstuk 4 wordt er in detail gekeken naar de toegevoegde waarde van mammografie als modaliteit in een screeningsprogramma waarin ook borst MRI gebruikt wordt. Uit eerder onderzoek is namelijk gebleken dat bij vrouwen met een hoge borstdensiteit mammografie een lagere sensitiviteit heeft[122]. Daarnaast zijn vrouwen die deelnemen aan hoog risico screeningsprogramma's relatief jong (wat een van de redenen is voor het aanwezig zijn van veel borstklierweefsel). In dit hoofdstuk is de invloed van mammografie op fout positieven (terugroepingen en biopsieën) en kanker detectie onderzocht in twee risico categorieën (*BRCA* mutatie draagsters en anderen) en meerdere leeftijdscategorieën. We concludeerden dat mammografie maar een beperkte waarde heeft als het gaat om kanker detectie in deze hoog risico populatie. In totaal werden 13 van de 125 kankers alleen gedetecteerd door mammografie. De resultaten laten zien dat mammografie alleen van waarde was in de oudere vrouwen-groep; in de groep vrouwen van onder de 40 jaar had mammografie geen toegevoegde waarde voor kanker detectie. Van de 13 vrouwen met tumoren alleen gedetecteerd door mammografie waren de meeste vrouwen ongeveer 50 jaar oud en bij veel (8 van de 13) van hen werd alleen een voorstadium (DCIS) van borstkanker ontdekt. Om bovenstaande redenen willen wij benadrukken dat de leeftijd waarop gestart wordt met mammografie kan worden verhoogd tot minimal 40 jaar, in tegenstelling tot de huidige richtlijn van 30 jaar. Dit zou vrouwen ook veel negatieve bijwerkingen van deze mammogrammen besparen, zoals de bijkomende blootstelling aan straling.

Omdat het primaire doel van screening het vroeg detecteren van kanker is, is het van groot belang dat kankers gedetecteerd worden op het juiste moment. Dat is de reden waarom hoofdstuk 5 geschreven is. In dit hoofdstuk hebben we onderzocht of borstkanker die gedetecteerd zijn in het uitgebreide screeningsprogramma (screeningsgedetecteerde kankers, interval kankers, incidenteel gevonden kankers in mastectomie specimen) al in een eerdere screeningsronde gedetecteerd hadden kunnen worden. In deze beoordeling van de voorafgaande MRI scans hebben we de onderverdeling gemaakt tussen zichtbaar, minimaal zichtbaar en onzichtbaar. Onze resultaten lieten zien dat bijna eenderde (31%) kon worden beschouwd als gemiste

kankers, omdat deze gescoord waren (in de re-evaluatie) als zichtbaar en verdacht op de voorafgaande MRI scan. Daarnaast werd 34% van de kankers gescoord als minimaal zichtbaar op deze voorafgaande MRI scan. De reden waarom deze kankers niet gedetecteerd werden in de klinische praktijk kan worden toegeschreven aan twee factoren: 1. het simpel niet opmerken van de borstkanker, of 2. het verkeerd interpreteren van de laesie. Dit moet verder worden uitgezocht in toekomstige studies.

Hoog risico screening wordt tegenwoordig uitgevoerd in een 'one-size-fits-all' strategie, waarin iedereen die in aanmerking komt voor deze intensieve screeningsprogrammas dezelfde onderzoeken ontvangt. Echter niet iedereen heeft hetzelfde borstkanker risico en dit risico kan ook veranderen tijdens het leven van een vrouw. Dit is een van de redenen waarom tegenwoordig veel onderzoek is gefocust op het personaliseren van screeningsprogramma's. In hoofdstuk 6 en 7 gaan we in op de mogelijkheden van gepersonaliseerde screeningsstrategieën voor vrouwen met een verhoogd risico op borstkanker. In meer detail hebben we onderzocht of twee potentiële eigenschappen van het normale borstweefsel gemeten op borst MRI gebruikt kunnen worden als risico voorspellende factoren. De hoeveelheid van fibroglandulair weefsel in de borst is een welbekende risico factor in vrouwen met een gemiddeld risico op borstkanker. Hoe meer fibroglandulair weefsel, hoe hoger het risico. Achtergrond aankleuring van het klierweefsel is de normale aankleuring van het klierweefsel na contrast toediening bij een borst MRI. King et al.[23] speculeerde dat deze factor mogelijk zou kunnen helpen bij het stratificeren van vrouwen in meer of minder intensieve screeningsregimes, omdat zij een hogere waarden van achtergrondaankleuring vonden in vrouwen met borstkanker vergeleken met vrouwen zonder borstkanker. Dit is de eerste aanzet geweest voor het schrijven van hoofdstuk 6, waar we de achtergrondaankleuring van borst MRI's (met borstkanker) hebben laten beoordelen door twee onafhankelijke waarnemers en we deze hebben gecorreleerd aan tumor karakteristieken. Onze resultaten lieten zien dat de door de patholoog bepaalde tumor graad negatief geassocieerd was met de waarden van achtergrondaankleuring, wat betekent dat bij vrouwen met een hoge achtergrondaankleuring de kansen hoger zijn om een kanker van lage graad met een betere prognose te detecteren dan bij vrouwen met een lagere achtergrondaankleuring. Dat betekent, als we de aanbevelingen van King et al. naleven, dit mogelijk zal leiden tot het detecteren van de meer traaggroeiende kankers en de meer agressieve kankers niet gedetecteerd zullen worden. Met dit hoofdstuk laten we zien dat er met zorg moet worden gekeken naar potentiële voorspellers voor patiënt stratificatie. In hoofdstuk 7 worden de potentiële risico factoren die uit borst MRI data herleid kunnen worden (de hoeveelheid borstklierweefsel en waarden van achtergrondaankleuring) in detail on-

derzocht. In dit hoofdstuk hebben we een risico predictie model ontworpen voor de ontwikkeling van borstkanker in de toekomst. We hebben een grote dataset gebruikt van 1533 borst MRI's. Om kwantitatieve en reproduceerbare waarden voor de hoeveelheid borstklierweefsel en achtergrondaankleuring te kunnen bepalen, hebben we gebruikt gemaakt van automatische segmentatie met behulp van een convolutioneel neuraal netwerk (diep-leren). De hoeveelheid borstklierweefsel en achtergrondaankleuring gemeten op de uitgangs-MRI zijn gebruikt in de berekeningen. De resultaten lieten zien dat de hoeveelheid borstklierweefsel noch achtergrondaankleuring voorspellend waren voor de toekomstige ontwikkeling van borstkanker in deze hoog risico populatie. Desalniettemin waren beiden gerelateerd aan een hoger risico op fout positieven in de eerste ronde van screening (uitgangs-MRI/ start van screening).

Een van de essenties van het personaliseren van het screeningsprotocol is dat het mogelijk de kosten-effectiviteit van deze screeningsprogramma's verbeterd. Er zijn echter ook andere manieren om de kosten-effectiviteit van borst MRI screening te verbeteren. Een relatief makkelijke manier is het reduceren van de tijd die het kost om een borst MRI onderzoek uit te voeren. Deze gereduceerde MRI protocollen worden vaak aangeduid als 'verkorte' borst MRI. Het huidige bewijs voor deze verkorte protocollen is gepresenteerd in hoofdstuk 8. De meeste studies rapporteren vergelijkbare sensitiviteitswaarden voor verkorte en volledige protocollen, hoewel dit soms ten koste gaat van een kleine vermindering in specificiteit.

Ultrasnelle borst MRI vormt een speciale tak binnen de verkorte protocollen, omdat het men nog steeds in de gelegenheid stelt om dynamische beoordelingen van contrast instroom uit te voeren en heeft daarom een zeker voordeel vergeleken met andere verkorte protocollen die alleen vertrouwen op de laesie morfologie. Helaas kan ultrasnelle borst MRI niet uitgevoerd worden op dezelfde spatiële resolutie als andere verkorte protocollen, omdat het data punten hergebruikt van verschillende tijdspunten. Dit proces wordt view-sharing wordt genoemd en leidt tot temporele vervaging. In hoofdstuk 9 onderzochten we de prestatie van een nieuwe sequentie voor borst MRI (CS-VIBE). Deze sequentie, die ook bestemd is voor ultrasnelle borst MRI, is gebasseerd op een techniek genaamd 'compressed sensing', dat in het kort beschreven kan worden als een techniek voor het efficiënt verkrijgen en reconstrueren van het MRI signaal door oplossingen te vinden voor ondergedefinieërde lineaire systemen[243]. Met het gebruik van deze sequentie (CS-VIBE) waren we in staat een volledig dynamisch protocol te verkrijgen binnen 2 minuten met een spatiële resolutie van 0.8 x 0.8 x 1.6 mm, zonder view-sharing te gebruiken. In deze 'proof-

of-concept' studie hebben we deze nieuwe sequentie vergeleken op beeldkwaliteit en laesie zichtbaarheid met een view-sharing (TWIST) en regulaire T1 gewogen opname (VIBE-DIXON). De resultaten lieten zien dat er geen significant verschil was tussen deze nieuwe sequentie en de dynamische evaluatie met TWIST en de spatiële resolutie was zo net zo goed als die van de regulaire VIBE-DIXON sequentie. CS-VIBE zou daarom mogelijk een interessant alternatief kunnen zijn voor de huidige sequenties in het borst MRI protocol. Dit brengt ons dus dichter bij "goedkope" borst MRI onderzoeken, zodat borst MRI in een grotere groep vrouwen gebruikt zou kunnen worden.

Samenvattend toont dit proefschrift de huidige prestatie van een uitgebreid screeningsprogramma voor vrouwen met een verhoogd risico op borstkanker in een academisch centrum. Verschillende zwakke punten zijn geïdentificeerd en potentiële punten voor verbetering zijn bediscussieerd, bijvoorbeeld het aantonen van het gebrek aan toegevoegde waarde van mammografie. Verder zijn potentiële risico factoren voor gepersonaliseerde screening onderzocht en bekritiseerd. Ten slotte werpt dit proefschrift een vluchtige blik op de mogelijkheden voor het verbeteren van de kosten-effectiviteit van borst MRI voor screening. Concluderend biedt dit proefschrift handvatten voor het verbeteren van borst MRI screening. Er zal echter nog verder onderzoek gedaan moeten worden om de resultaten genoemd in dit proefschrift te valideren.

Publications

Papers in (inter-) national journals

Gubern-Merida A, **Vreemann S**, Marti R, et al. Automated detection of breast cancer in false-negative screening MRI studies from women at increased risk. *Eur J Radiol*. 2016;85(2):472-479.

Dalmis MU, Gubern-Merida A, **Vreemann S**, Karssemeijer N, Mann R, Platel B. A computer-aided diagnosis system for breast DCE-MRI at high spatiotemporal resolution. *Med Phys*. 2016;43(1):84.

Mertzanidou T, Hipwell JH, Reis S, Hawkes B, Ehteshami Bejnordi B, Dalmis MU, **Vreemann S**, Platel B, van der Laak J, Karssemeijer N, Hermsen M, Bult P, Mann RM. 3D volume reconstruction from serial breast specimen radiographs for mapping between histology and 3D whole specimen imaging. *Med Phys*. 2017;44(3):935-948.

Vreemann S, Rodriguez-Ruiz A, Nickel D, Heacock L, Appelman L, van Zelst J, Karssemeijer N, Weiland E, Maas M, Moy L, Kiefer B, Mann RM. Compressed Sensing for Breast MRI: Resolving the Trade-Off Between Spatial and Temporal Resolution. *Invest Radiol*. 2017;52:574-582.

van Zelst JCM, Mus RDM, Woldringh G, Rutten M, Bult P, **Vreemann S**, de Jong M, Karssemeijer N, Hoogerbrugge N, Mann RM. Surveillance of Women with the BRCA1 or BRCA2 Mutation by Using Biannual Automated Breast US, MR Imaging, and Mammography. *Radiology*. 2017;285(2):376-388.

Vreemann S, Gubern-Merida A, Schlooz-Vries MS, et al. Influence of Risk Category and Screening Round on the Performance of an MR Imaging and Mammography Screening Program in Carriers of the BRCA Mutation and Other Women at Increased Risk. *Radiology*. 2018;286(2):443-451.

Rodriguez-Ruiz A, Feng SSJ, van Zelst J, **Vreemann S**, Mann JR, D'Orsi CJ, Sechopoulos I. Improvements of an objective model of compressed breasts undergoing mammography: Generation and characterization of breast shapes. *Med Phys*. 2017;44(6): 2161-2172.

Mann RM and **Vreemann S**. CS-VIBE - a breakthrough in Ultrafast Dynamic Breast MRI. *MAGNETOM Flash*. (68) 2/2017.

Karssemeijer N, Teuwen J, **Vreemann S**, Mann RM. Deep learning-technieken voor detectie en diagnose van borstkanker. *Memorad*. (22) 2/2017.

Mann RM, van Zelst JCM, **Vreemann S**, Mus RDM. Is ultrafast or abbreviated breast MRI ready for prime time? *Current reports in breast cancer research*. 2017. *Accepted*

Vreemann S, Gubern-Merida A, Borelli C, Bult P, Karssemeijer N, Mann RM. The correlation of background parenchymal enhancement in the contralateral breast with patient and tumor characteristics of MRI-screen detected breast cancers. *PLOS ONE*. 2018;13(1): e0191399.

Rodriguez-Ruiz A, Teuwen J, **Vreemann S**, Bouwman RW, van Engen RE, Karssemeijer N, Mann RM, Gubern-Merida A, Sechopoulos I. New reconstruction algorithm for digital breast tomosynthesis: better image quality for humans and computers. *Acta Radiol*. 2018 Sep;59(9):1051-1059.

Dalmis MU, **Vreemann S**, Kooi T, Mann RM, Karssemeijer N, Mann R, Gubern-Merida A. Fully automated detection of breast cancer in screening MRI using convolutional neural networks. *J Med Imaging*. 2018;5(1),014502.

Vreemann S, Gubern-Merida A, Lardenoije S, Bult P, Karssemeijer N, Pinker K, Mann RM. The frequency of missed breast cancers in women participating in a high-risk MRI screening program. *Breast Cancer Res Treat*. 2018 Jun;169(2):323-331.

van Zelst JCM, **Vreemann S**, Witt HJ, et al. Multireader study on the diagnostic accuracy of ultrafast breast Magnetic Resonance Imaging for breast cancer screening. *Invest Radiol*. 2018 Jun 25. *Epub ahead of print*

Vreemann S, van Zelst JCM, Schlooz-Vries M, Bult P, Hoogerbrugge N, Karssemeijer N, Gubern-Merida A, Mann RM. The added value of mammography in different age-groups of women with and without BRCA mutation screened with breast MRI. *Breast Cancer Res*. 2018 Aug 3;20(1):84

Vreemann S, Dalmis MU, Bult P, Karssemeijer N, Broeders MJM, Gubern-Merida A, Mann RM. Amount of fibroglandular tissue and background parenchymal enhancement in relation to risk for breast cancer and false positives in a breast MRI screening program. 2018. *Submitted*

Abstracts in conference proceedings

Vreemann S, Gubern-Merida A, Lardenoije S, Platel B, Karssemeijer N, Mann RM. A critical audit of a breast MRI screening programme for intermediate and high risk patients in clinical practice. In: *European Congress of Radiology*, 2015. *Oral Presentation*

Gubern-Merida A, **Vreemann S**, Marti R, Melendez J, Lardenoije S, Mann RM, Karssemeijer N. Automated detection of breast cancer as an aid in the interpretation of screening MRI. In: *European Congress of Radiology*, 2015. *Oral Presentation*

Dalmis MU, Gubern-Merida A, **Vreemann S**, Platel B, Mann RM, Karssemeijer N. Is late-phase information necessary for dynamic evaluation of breast cancer? In: *European Congress of Radiology*, 2015. *Oral Presentation*

Vreemann S, Gubern-Merida A, Lardenoije S, Platel B, Karssemeijer N, Mann RM. Gemiste borstkanker op MRI. In: *Radiologendagen*, 2015. *Invited talk*

Vreemann S, Gubern-Merida A, Lardenoije S, Platel B, Karssemeijer N, Mann RM. Longitudinal results of a breast MRI screening program for patients at high and intermediate risk; does BRCA status matter? In: *Radiological Society of North America*, 2015. *Oral Presentation*

Vreemann S, Gubern-Merida A, Lardenoije S, Platel B, Karssemeijer N, Mann RM. Breast cancers not detected by MRI in a high and intermediate risk screening program. In: *Radiological Society of North America*, 2015. *Oral Presentation*

Dalmis MU, Gubern-Merida A, **Vreemann S**, Mann RM, Karssemeijer N, Platel B. Early Phase Contrast Enhancement Dynamics of Breast Lesions of Different Molecular Subtypes Characterized by a Computer-Aided-Diagnosis System. In: *Radiological Society of North America*, 2015. *Oral Presentation*

Gubern-Merida A, **Vreemann S**, Marti R, Mendelez J, Mann RM, Platel B, Karssemeijer N. Automated detection of mass-like, non-mass-like and focus breast cancer lesions visible in false-negative screening DCE-MRI. In: *Radiological Society of North America*, 2015. *Oral Presentation*

Vreemann S, Gubern-Merida A, Borelli C, Karssemeijer N, Mann RM. Background parenchymal enhancement as a predictor of breast cancer grade: a pilot study. In:

European Congress of Radiology, 2016. *Oral Presentation*

Vreemann S, Gubern-Merida A, Lardenoije S, Karssemijer N, Mann RM. Prognostic factors of interval carcinomas occurring in an intermediate and high risk breast cancer screening program. In: *European Congress of Radiology*, 2016. *Oral Presentation*

Vreemann S, Gubern-Merida A, Lardenoije S, Karssemeijer N, Mann RM. Differences between cancers detected in prophylactic mastectomy specimen, screen detected cancers and true interval cancers in women participating in an intermediate and high risk screening program. In: *European Breast Cancer Conference*, 2016. *Poster Presentation*

Dalmis MU, Gubern-Merida A, Borelli C, **Vreemann S**, Mann RM, Karssemeijer. A fully automated system for quantification of background parenchymal enhancement in breast DCE-MRI. *SPIE*. 2016.

Vreemann S, Gubern-Merida A, Lardenoije S, Karssemeijer N, Mann RM. The performance of MRI screening in the detection of breast cancer in an intermediate and high risk screening program. In: *International Society for Magnetic Resonance in Medicine*, 2016. *Oral Presentation*

Honroth T, Vicari M, **Vreemann S**, Mann RM, Gunther M. 3D Arterial Spin Labeling in breast cancer: a case study. In: *International Society for Magnetic Resonance in Medicine*, 2016. *Poster Presentation*

Vreemann S, Gubern-Merida A, van Zelst J. Karssemeijer N, Mann RM. The complimentary value of mammography and MRI in the detection of in-situ breast cancer in an intermediate and high risk breast cancer screening program. In: *European Society of Breast Imaging*, 2016. *Poster Presentation*

Vreemann S, Gubern-Merida A, Lardenoije S, Karssemeijer N, Mann RM. Differences in prognostic characteristics in breast cancers detected in an intermediate and high risk breast cancer screening program, stratified by detection mode. In: *European Society of Breast Imaging*, 2016. *Poster Presentation*

Vreemann S. Evaluation of breast MRI as a screening tool. In: *Advanced Breast Imaging Course*, 2016. *Invited talk*

Dalmis MU, Gubern-Merida A, Borelli C, **Vreemann S**, Mann RM, Karssemeijer N. A Fully Automated System for Quantification of Background Parenchymal Enhancement in Breast DCE-MRI. In: *International Society for Optics and Photonics*, 2016. *Oral Presentation*

Saleem MA, Vicari M, Kompan IN, Mann R, **Vreemann S**. Sparse Parallel Acceleration for full 3D Golden Angle Radial Whole Breast DCE MR. In: *European Society for Magnetic Resonance in Medicine and Biology*, 2016. *Poster Presentation*

Vreemann S, van Zelst J, Gubern-Merida A, Karssemeijer N, Mann RM. The added value of mammography in an intermediate and high risk breast cancer screening program. In: *Radiological Society of North America*, 2016. *Oral Presentation*

Vreemann S, Rodriguiz-Ruiz A, Nickel D, Heacock L, Appelman L, van Zelst J, Karssemeijer N, Weiland E, Maas M, Moy L, Kiefer B, Mann RM. Advanced ultrafast Dynamic Contrast Enhanced breast MRI with Compressed Sensing. In: *International Society for Magnetic Resonance in Medicine*, 2017. *Poster Presentation*

Vreemann S, Dalmis MU, Karssemeijer N, Broeders M, Gubern-Merida A, Mann RM. Is the Amount of Fibroglandular Breast Tissue and Background Parenchymal Enhancement Risk Predictors for Breast Cancer Development and False Positive Results in a Breast MRI Screening Program? In: *Radiological Society of North America*, 2017. *Oral Presentation*

Gubern-Merida A, Wichakam I, Rodriguiz-Ruiz A, **Vreemann S**, Bouwman R, van Engen RE. New Reconstruction Algorithm for Breast Tomosynthesis: Better Image Quality for Humans and Computers. In: *Radiological Society of North America*, 2017. *Poster Presentation*

Bibliography

[1] van der Waal D., Verbeek A. L., den Heeten G. J., Ripping T. M., Tjan-Heijnen V. C., and Broeders M. J. Breast cancer diagnosis and death in the Netherlands: a changing burden. *Eur J Public Health*, 25(2):320–4, 2015.

[2] Rijksinstituut voor Volksgezondheid en Milieu. Breast cancer in the Netherlands. `http://www.rivm.nl/en/Topics/B/Breast_cancer_screening_programme/Breast_cancer_in_the_Netherlands`, 2017. Accessed: 2017-01-29.

[3] Li C. I., Anderson B. O., Daling J. R., and Moe R. E. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA*, 289(11):1421–4, 2003.

[4] Barroso-Sousa R. and Metzger-Filho O. Differences between invasive lobular and invasive ductal carcinoma of the breast: results and therapeutic implications. *Ther Adv Med Oncol*, 8(4): 261–266, 2016.

[5] Dieci M., Orvieto E., Dominici M., Conte P., and Guarneri V. Rare breast cancer subtypes: histological, molecular, and clinical peculiarities. *Oncologist*, 19:805–813, 2014.

[6] Holland R. and Hendriks J. H. Microcalcifications associated with ductal carcinoma in situ: mammographic-pathologic correlation. *Seminars in Diagnostic Pathology*, 11(3):181–92, 1994.

[7] Freer P. E. Mammographic breast density: impact on breast cancer risk and implications for screening. *Radiographics*, 35(2):302–15, 2015.

[8] Ehteshami Bejnordi B. *Histopathological diagnosis of breast cancer using machine learning*. PhD thesis, Radboud University Medical Center, 2017.

[9] Emaus M. J., Bakker M. F., Peeters P. H., Loo C. E., Mann R. M., de Jong M. D., Bisschops R. H., Veltman J., Duvivier K. M., Lobbes M. B., Pijnappel R. M., Karssemeijer N., de Koning H. J., van den Bosch M. A., Monninkhof E. M., Mali W. P., Veldhuis W. B., and van Gils C. H. Mr Imaging as an additional screening modality for the detection of breast cancer in women aged 50-75 years with extremely dense breasts: The DENSE Trial Study design. *Radiology*, 277(2): 527–37, 2015.

[10] Burnside E. S., DeMartini W. B., schrager S., Trentham-Dietz A., Hampton J., Shafer C., and Wilke L. Trade-offs of risk-based versus age-based mammography screening in women aged ¡ 50 years. In *Proc. RSNA*, Chicago, IL, 2016.

[11] Mann R. M., Kuhl C. K., Kinkel K., and Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. *Eur Radiol*, 18(7):1307–18, 2008.

[12] Saslow D., Boetes C., Burke W., Harms S., Leach M. O., Lehman C. D., Morris E., Pisano E., Schnall M., Sener S., Smith R. A., Warner E., Yaffe M., Andrews K. S., and Russell C. A. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*, 57(2):75–89, 2007.

[13] Bushberg J., Seibert J., Leidholdt E., and Boone J. *The Essential Physics of Medical Imaging*. 3 edition, 2012.

[14] Kriege M., Brekelmans C. T., Boetes C., Besnard P. E., Zonderland H. M., Obdeijn I. M., Manoliu R. A., Kok T., Peterse H., Tilanus-Linthorst M. M., Muller S. H., Meijer S., Oosterwijk J. C., Beex L. V., Tollenaar R. A., de Koning H. J., Rutgers E. J., and Klijn J. G. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *N Engl J Med*, 351(5):427–37, 2004.

[15] Riedl C. C., Luft N., Bernhart C., Weber M., Bernathova M., Tea M. K., Rudas M., Singer C. F., and Helbich T. H. Triple-modality screening trial for familial breast cancer underlines the importance of magnetic resonance imaging and questions the role of mammography and ultrasound regardless of patient mutation status, age, and breast density. *J Clin Oncol*, 33(10):1128–35, 2015.

[16] Harvey S. C., Di Carlo P. A., Lee B., Obadina E., Sippo D., and Mullen L. An abbreviated protocol for high-risk screening breast MRI saves time and resources. *J Am Coll Radiol*, 13(4): 374–80, 2016.

[17] Moschetta M., Telegrafo M., Rella L., Stabile Ianora A. A., and Angelelli G. Abbreviated combined MR protocol: A new faster strategy for characterizing breast lesions. *Clin Breast Cancer*, 16(3):207–11, 2016.

[18] Kuhl C. K., Schrading S., Bieling H. B., Wardelmann E., Leutner C. C., Koenig R., Kuhn W., and Schild H. H. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet*, 370(9586):485–92, 2007.

[19] Edwards S. D., Lipson J. A., Ikeda D. M., and Lee J. M. Updates and revisions to the BI-RADS magnetic resonance imaging lexicon. *Magn Reson Imaging Clin N Am*, 21(3):483–93, 2013.

[20] NABON. Breast Cancer Guideline. 2012.

[21] Gubern-Merida A. *Automated analysis of magnetic resonance imaging of the breast*. PhD thesis, Radboud University Medical Center, 2015.

[22] Rao A. A., Feneis J., Lalonde C., and Ojeda-Fournier H. A pictorial review of changes in the BI-RADS Fifth Edition. *Radiographics*, 36(3):623–39, 2016.

[23] King V., Brooks J. D., Bernstein J. L., Reiner A. S., Pike M. C., and Morris E. A. Background parenchymal enhancement at breast MR imaging and breast cancer risk. *Radiology*, 260(1):50–60, 2011.

[24] Dontchos B. N., Rahbar H., Partridge S. C., Korde L. A., Lam D. L., Scheel J. R., Peacock S., and Lehman C. D. Are qualitative assessments of background parenchymal enhancement, amount of fibroglandular tissue on MR images, and mammographic density associated with breast cancer risk? *Radiology*, 276(2):371–80, 2015.

[25] Lang K., Andersson I., Rosso A., Tingberg A., Timberg P., and Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. *Eur Radiol*, 26(1): 184–90, 2016.

[26] Fornvik D., Andersson I., Svahn T., Timberg P., Zackrisson S., and Tingberg A. The effect of reduced breast compression in breast tomosynthesis: human observer study using clinical cases. *Radiat Prot Dosimetry*, 139(1-3):118–23, 2010.

[27] Svahn T. M., Houssami N., Sechopoulos I., and Mattsson S. Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. *Breast*, 24(2):93–9, 2015.

[28] Spangler M. L., Zuley M. L., Sumkin J. H., Abrams G., Ganott M. A., Hakim C., Perrin R., Chough D. M., Shah R., and Gur D. Detection and classification of calcifications on digital breast tomosynthesis and 2D digital mammography: a comparison. *AJR Am J Roentgenol*, 196

(2):320–4, 2011.

[29] Kopans D., Gavenonis S., Halpern E., and Moore R. Calcifications in the breast and digital breast tomosynthesis. *Breast J*, 17(6):638–44, 2011.

[30] Clauser P., Nagl G., Helbich T. H., Pinker-Domenig K., Weber M., Kapetas P., Bernathova M., and Baltzer P. A. T. Diagnostic performance of digital breast tomosynthesis with a wide scan angle compared to full-field digital mammography for the detection and characterization of microcalcifications. *Eur J Radiol*, 85(12):2161–2168, 2016.

[31] Brem R. F., Tabar L., Duffy S. W., Inciardi M. F., Guingrich J. A., Hashimoto B. E., Lander M. R., Lapidus R. L., Peterson M. K., Rapelyea J. A., Roux S., Schilling K. J., Shah B. A., Torrente J., Wynn R. T., and Miller D. P. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight study. *Radiology*, 274(3):663–73, 2015.

[32] van Zelst J. C. M., Mus R. D. M., Woldringh G., Rutten M., Bult P., Vreemann S., de Jong M., Karssemeijer N., Hoogerbrugge N., and Mann R. M. Surveillance of women with the BRCA1 or BRCA2 mutation by using Biannual Automated Breast US, MR imaging, and Mammography. *Radiology*, 285(2):376–388, 2017.

[33] Rosenberg R. D., Yankaskas B. C., Abraham L. A., Sickles E. A., Lehman C. D., Geller B. M., Carney P. A., Kerlikowske K., Buist D. S., Weaver D. L., Barlow W. E., and Ballard-Barbash R. Performance benchmarks for screening mammography. *Radiology*, 241(1):55–66, 2006.

[34] Perry N., Broeders M., de Wolf, C. T. S., Holland R., and von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition – summary document. *Ann Oncol*, 19(4):614–622, 2008.

[35] Heywang S. H., Hahn D., Schmidt H., Krischke I., Eiermann W., Bassermann R., and Lissner J. MR imaging of the breast using gadolinium-DTPA. *J Comput Assist Tomogr*, 10(2):199–204, 1986.

[36] Heywang S. H., Hilbertz T., Pruss E., Wolf A., Permanetter W., Eiermann W., and Lissner J. [dynamic contrast medium studies with flash sequences in nuclear magnetic resonance tomography of the breast]. *Digitale Bilddiagnostik*, 8(1):7–13, 1988.

[37] Leach M. O., Boggis C. R., Dixon A. K., Easton D. F., Eeles R. A., Evans D. G., Gilbert F. J., Griebsch I., Hoff R. J., Kessar P., Lakhani S. R., Moss S. M., Nerurkar A., Padhani A. R., Pointon L. J., Thompson D., and Warren R. M. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). *Lancet*, 365(9473):1769–78, 2005.

[38] Kuhl C. K., Schrading S., Leutner C. C., Morakkabati-Spitz N., Wardelmann E., Fimmers R., Kuhn W., and Schild H. H. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol*, 23(33):8469–76, 2005.

[39] Rijnsburger A. J., Obdeijn I. M., Kaas R., Tilanus-Linthorst M. M., Boetes C., Loo C. E., Wasser M. N., Bergers E., Kok T., Muller S. H., Peterse H., Tollenaar R. A., Hoogerbrugge N., Meijer S., Bartels C. C., Seynaeve C., Hooning M. J., Kriege M., Schmitz P. I., Oosterwijk J. C., de Koning H. J., Rutgers E. J., and Klijn J. G. BRCA1-associated breast cancers present differently from BRCA2-associated and familial cases: long-term follow-up of the Dutch MRISC screening study. *J Clin Oncol*, 28(36):5265–73, 2010.

[40] Trop I., Lalonde L., Mayrand M. H., David J., Larouche N., and Provencher D. Multimodality breast cancer screening in women with a familial or genetic predisposition. *Curr Oncol*, 17(3): 28–36, 2010.

[41] Sardanelli F., Podo F., Santoro F., Manoukian S., Bergonzi S., Trecate G., Vergnaghi D., Federico M., Cortesi L., Corcione S., Morassut S., Di Maggio C., Cilotti A., Martincich L., Calabrese M., Zuiani C., Preda L., Bonanni B., Carbonaro L. A., Contegiacomo A., Panizza P., Di Cesare E., Savarese A., Crecco M., Turchetti D., Tonutti M., Belli P., and Maschio A. D. Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. *Invest Radiol*, 46(2):94–105, 2011.

[42] Passaperuma K., Warner E., Causer P. A., Hill K. A., Messner S., Wong J. W., Jong R. A., Wright F. C., Yaffe M. J., Ramsay E. A., Balasingham S., Verity L., Eisen A., Curpen B., Shumak R., Plewes D. B., and Narod S. A. Long-term results of screening with magnetic resonance imaging in women with BRCA mutations. *Br J Cancer*, 107(1):24–30, 2012.

[43] Gillman J., Toth H. K., and Moy L. The role of dynamic contrast-enhanced screening breast MRI in populations at increased risk for breast cancer. *Womens Health (Lond)*, 10(6):609–22, 2014.

[44] Sardanelli F., Boetes C., Borisch B., Decker T., Federico M., Gilbert F. J., Helbich T., Heywang-Kobrunner S. H., Kaiser W. A., Kerin M. J., Mansel R. E., Marotti L., Martincich L., Mauriac L., Meijers-Heijboer H., Orecchia R., Panizza P., Ponti A., Purushotham A. D., Regitnig P., Del Turco M. R., Thibault F., and Wilson R. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. *Eur J Cancer*, 46(8):1296–316, 2010.

[45] DeMartini W. B. and Rahbar H. Breast magnetic resonance imaging technique at 1.5 T and 3 T: requirements for quality imaging and American College of Radiology Accreditation. *Magn Reson Imaging Clin N Am*, 21(3):475–82, 2013.

[46] Orel S. G. MR imaging of the breast. *Radiol Clin North Am*, 38(4):899–913, 2000.

[47] Lourenco A. P., Donegan L., Khalil H., and Mainiero M. B. Improving outcomes of screening breast MRI with practice evolution: initial clinical experience with 3T compared to 1.5T. *J Magn Reson Imaging*, 39(3):535–9, 2014.

[48] Dhillon G. S., Bell N., Ginat D. T., Levit A., Destounis S., and O'Connell A. Breast MR imaging: What the radiologist needs to know. *J Clin Imaging Sci*, 1:48, 2011.

[49] Gaillard F. and Ballinger J. MRI pulse sequences. `https://radiopaedia.org/`, 2017. Accessed: 2017-02-20.

[50] Kuhl C. K. and Schild H. H. Dynamic image interpretation of MRI of the breast. *J Magn Reson Imaging*, 12(6):965–74, 2000.

[51] D'Orsi C., Sickles E., Mendelson E., and Morris E. *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*. American college of Radiology, 2013.

[52] Mango V. L., Morris E. A., David Dershaw D., Abramson A., Fry C., Moskowitz C. S., Hughes M., Kaplan J., and Jochelson M. S. Abbreviated protocol for breast MRI: Are multiple sequences needed for cancer detection? *Eur J Radiol*, 84(1):65–70, 2015.

[53] Grimm L. J., Soo M. S., Yoon S., Kim C., Ghate S. V., and Johnson K. S. Abbreviated screening protocol for breast MRI: a feasibility study. *Acad Radiol*, 22(9):1157–62, 2015.

[54] Heacock L., Melsaether A. N., Heller S. L., Gao Y., Pysarenko K. M., Babb J. S., Kim S. G., and Moy L. Evaluation of a known breast cancer using an abbreviated breast MRI protocol: Correlation of imaging characteristics and pathology with lesion detection and conspicuity. *Eur J Radiol*, 85(4):815–23, 2016.

[55] Mann R. M., Mus R. D., van Zelst J., Geppert C., Karssemeijer N., and Platel B. A novel approach to contrast-enhanced breast magnetic resonance imaging for screening: high-resolution ultrafast dynamic imaging. *Invest Radiol*, 49(9):579–85, 2014.

[56] Elster A. Questions and answers in MRI. `http://www.mriquestions.com/index.html`, 2017. Accessed: 2017-08-08.

[57] Kuhl C. K., Klaschik S., Mielcarek P., Gieseke J., Wardelmann E., and Schild H. H. Do T2-weighted pulse sequences help with the differential diagnosis of enhancing lesions in dynamic breast MRI? *J Magn Reson Imaging*, 9(2):187–96, 1999.

[58] Hendrick R. E. Breast MRI: Using physics to maximize its sensitivity and specificity to breast cancer. *Med Phys*, 31(6):1737–1737, 2004.

[59] Hagmann P., Jonasson L., Maeder P., Thiran J. P., Wedeen V. J., and Meuli R. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics*, 26 Suppl 1:S205–23, 2006.

[60] Dorrius M. D., Dijkstra H., Oudkerk M., and Sijens P. E. Effect of b value and pre-admission of contrast on diagnostic accuracy of 1.5-T breast DWI: a systematic review and meta-analysis. *Eur Radiol*, 24(11):2835–47, 2014.

[61] Partridge S. C. and McDonald E. S. Diffusion weighted magnetic resonance imaging of the breast: protocol optimization, interpretation, and clinical applications. *Magn Reson Imaging Clin N Am*, 21(3):601–24, 2013.

[62] Pinker K., Bickel H., Helbich T. H., Gruber S., Dubsky P., Pluschnig U., Rudas M., Bago-Horvath Z., Weber M., Trattnig S., and Bogner W. Combined contrast-enhanced magnetic resonance and diffusion-weighted imaging reading adapted to the "Breast Imaging Reporting and Data System" for multiparametric 3-T imaging of breast lesions. *Eur Radiol*, 23(7):1791–802, 2013.

[63] Bickel H., Pinker-Domenig K., Bogner W., Spick C., Bago-Horvath Z., Weber M., Helbich T., and Baltzer P. Quantitative Apparent Diffusion Coefficient as a noninvasive imaging biomarker for the differentiation of invasive breast cancer and ductal carcinoma in situ. *Invest Radiol*, 50(2): 95–100, 2015.

[64] Kul S., Cansu A., Alhan E., Dinc H., Gunes G., and Reis A. Contribution of Diffusion-Weighted Imaging to dynamic contrast-enhanced MRI in the characterization of breast tumors. *AJR Am J Roentgenol*, 196(1):210–217, 2011.

[65] Bansal R., Shah V., and Aggarwal B. Qualitative and quantitative diffusion-weighted imaging of the breast at 3t - a useful adjunct to contrast-enhanced MRI in characterization of breast lesions. *Indian J Radiol Imaging*, 25(4):397–403, 2015.

[66] Chavhan G. B., Babyn P. S., Thomas B., Shroff M. M., and Haacke E. M. Principles, techniques, and applications of T2*-based MR imaging and its special applications. *Radiographics*, 29(5): 1433–49, 2009.

[67] Kvistad K. A., Rydland J., Vainio J., Smethurst H. B., Lundgren S., Fjosne H. E., and Haraldseth

O. Breast lesions: evaluation with dynamic contrast-enhanced T1-weighted MR imaging and with T2*-weighted first-pass perfusion MR imaging. *Radiology*, 216(2):545–53, 2000.

[68] Sardanelli F., Fausto A., and Podo F. MR spectroscopy of the breast. *Radiol Med*, 113(1):56–64, 2008.

[69] Sardanelli F., Fausto A., Di Leo G., de Nijs R., Vorbuchner M., and Podo F. In vivo proton MR spectroscopy of the breast using the total choline peak integral as a marker of malignancy. *AJR Am J Roentgenol*, 192(6):1608–17, 2009.

[70] Montemezzi S., Cavedon C., Camera L., Meliado G., Caumo F., Baglio I., and Sardanelli F. 1H-MR spectroscopy of suspicious breast mass lesions at 3T: a clinical experience. *Radiol Med*, 122 (3):161–170, 2017.

[71] Sardanelli F., Carbonaro L. A., Montemezzi S., Cavedon C., and Trimboli R. M. Clinical breast MR using MRS or DWI: Who is the winner? *Front Oncol*, 6:217, 2016.

[72] Bernstein M. A., Huston J., and Ward H. A. Imaging artifacts at 3.0T. *J Magn Reson Imaging*, 24 (4):735–746, 2006.

[73] Czervionke L. F., Czervionke J. M., Daniels D. L., and Haughton V. M. Characteristic features of MR truncation artifacts. *AJR Am J Roentgenol*, 151(6):1219–28, 1988.

[74] Kuhl C. K., Jost P., Morakkabati N., Zivanovic O., Schild H. H., and Gieseke J. Contrast-enhanced MR imaging of the breast at 3.0 and 1.5 T in the same patients: initial experience. *Radiology*, 239(3):666–76, 2006.

[75] Kuhl C., Weigel S., Schrading S., Arand B., Bieling H., Konig R., Tombach B., Leutner C., Rieber-Brambs A., Nordhoff D., Heindel W., Reiser M., and Schild H. H. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *J Clin Oncol*, 28(9):1450–7, 2010.

[76] Kuhl C. K., Schrading S., Strobel K., Schild H. H., Hilgers R. D., and Bieling H. B. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection-a novel approach to breast cancer screening with MRI. *J Clin Oncol*, 32(22): 2304–10, 2014.

[77] Pineda F. D., Medved M., Wang S., Fan X., Schacht D. V., Sennett C., Oto A., Newstead G. M., Abe H., and Karczmar G. S. Ultrafast bilateral DCE-MRI of the breast with conventional Fourier sampling: Preliminary evaluation of semi-quantitative analysis. *Acad Radiol*, 2016.

[78] Herrmann K. H., Baltzer P. A., Dietzel M., Krumbein I., Geppert C., Kaiser W. A., and Reichenbach J. R. Resolving arterial phase and temporal enhancement characteristics in DCE MRM at high spatial resolution with TWIST acquisition. *J Magn Reson Imaging*, 34(4):973–982, 2011.

[79] Tudorica L. A., Oh K. Y., Roy N., Kettler M. D., Chen Y. Y., Hemmingson S. L., Afzal A., Grinstead J. W., Laub G., Li X., and Huang W. A feasible high spatiotemporal resolution breast DCE-MRI protocol for clinical settings. *Magn Reson Imaging*, 30(9):1257–1267, 2012.

[80] Le Y., Kroeker R., Kipfer H. D., and Lin C. Development and evaluation of TWIST Dixon for dynamic contrast-enhanced (DCE) MRI with improved acquisition efficiency and fat suppression. *J Magn Reson Imaging*, 36(2):483–491, 2012.

[81] van Zelst J., Gubern-Merida A., Vreemann S., Veltman J., Dorrius M., Loo C., Hoof K. D. v.,

Karssemeijer N., and Mann R. High risk breast cancer screening with an ultrafast high spatiotemporal resolution MRI sequence; less costly and as reliable as a full diagnostic MRI protocol. In *Proc. RSNA*, Chicago, IL, 2016.

[82] Feng L., Grimm R., Block K. T., Chandarana H., Kim S., Xu J., Axel L., Sodickson D. K., and Otazo R. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn Reson Med*, 72(3):707–17, 2014.

[83] Heacock L., Gao Y., Heller S. L., Melsaether A. N., Babb J. S., Block T. K., Otazo R., Kim S. G., and Moy L. Comparison of conventional DCE-MRI and a novel golden-angle radial multicoil compressed sensing method for the evaluation of breast lesion conspicuity. *J Magn Reson Imaging*, 45(6):1746–1752, 2017.

[84] Kanda T., Fukusato T., Matsuda M., Toyoda K., Oba H., Kotoku J., Haruyama T., Kitajima K., and Furui S. Gadolinium-based contrast agent accumulates in the brain even in subjects without severe renal dysfunction: Evaluation of autopsy brain specimens with inductively coupled plasma mass spectroscopy. *Radiology*, 276(1):228–32, 2015.

[85] Radbruch A., Weberling L. D., Kieslich P. J., Eidel O., Burth S., Kickingereder P., Heiland S., Wick W., Schlemmer H. P., and Bendszus M. Gadolinium retention in the dentate nucleus and globus pallidus is dependent on the class of contrast agent. *Radiology*, 275(3):783–91, 2015.

[86] McDonald R. J., McDonald J. S., Kallmes D. F., Jentoft M. E., Murray D. L., Thielen K. R., Williamson E. E., and Eckel L. J. Intracranial gadolinium deposition after contrast-enhanced MR Imaging. *Radiology*, 275(3):772–82, 2015.

[87] Mainiero M. B., Lourenco A., Mahoney M. C., Newell M. S., Bailey L., Barke L. D., D'Orsi C., Harvey J. A., Hayes M. K., Huynh P. T., Jokich P. M., Lee S. J., Lehman C. D., Mankoff D. A., Nepute J. A., Patel S. B., Reynolds H. E., Sutherland M. L., and Haffty B. G. ACR appropriateness criteria breast cancer screening. *J Am Coll Radiol*, 13(11S):R45–R49, 2016.

[88] Bickelhaupt S., Tesdorff J., Laun F. B., Kuder T. A., Lederer W., Teiner S., Maier-Hein K., Daniel H., Stieber A., Delorme S., and Schlemmer H. P. Independent value of image fusion in unenhanced breast MRI using diffusion-weighted and morphological T2-weighted images for lesion characterization in patients with recently detected BI-RADS 4/5 X-ray mammography findings. *Eur Radiol*, 27(2):562–569, 2016.

[89] Furman-Haran E., Eyal E., Shapiro-Feinberg M., Nissan N., Grobgeld D., Weisenberg N., and Degani H. Advantages and drawbacks of breast DTI. *Eur J Radiol*, 81 Suppl 1:S45–7, 2012.

[90] Liu C., Liang C., Liu Z., Zhang S., and Huang B. Intravoxel incoherent motion (IVIM) in evaluation of breast lesions: comparison with conventional DWI. *Eur J Radiol*, 82(12):e782–9, 2013.

[91] Buchbender S., Obenauer S., Mohrmann S., Martirosian P., Buchbender C., Miese F. R., Wittsack H. J., Miekley M., Antoch G., and Lanzman R. S. Arterial spin labelling perfusion MRI of breast cancer using FAIR TrueFISP: initial results. *Clin Radiol*, 68(3):e123–7, 2013.

[92] Zhu D. C. and Buonocore M. H. Breast tissue differentiation using arterial spin tagging. *Magn Reson Med*, 50(5):966–75, 2003.

[93] Kawashima M., Katada Y., Shukuya T., Kojima M., and Nozaki M. MR perfusion imaging using the arterial spin labeling technique for breast cancer. *J Magn Reson Imaging*, 35(2):436–40, 2012.

[94] Podo F., Paris L., Cecchetti S., Spadaro F., Abalsamo L., Ramoni C., Ricci A., Pisanu M. E., Sardanelli F., Canese R., and Iorio E. Activation of Phosphatidylcholine-specific Phospholipase C in breast and ovarian cancer: Impact on MRS-detected Choline metabolic profile and perspectives for targeted therapy. *Front Oncol*, 6:171, 2016.

[95] Schmitt B., Trattnig S., and Schlemmer H. P. CEST-imaging: A new contrast in MR-mammography by means of chemical exchange saturation transfer. *Eur J Radiol*, 81 Suppl 1: S144–6, 2012.

[96] Chen S. and Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J Clin Oncol*, 25 (11):1329–33, 2007.

[97] Narod S. A. Screening of women at high familial risk of breast cancer. *Womens Health (Lond)*, 6 (4):509–10, 2010.

[98] Brekelmans C. T., Seynaeve C., Bartels C. C., Tilanus-Linthorst M. M., Meijers-Heijboer E. J., Crepin C. M., van Geel A. A., Menke M., Verhoog L. C., van den Ouweland A., Obdeijn I. M., and Klijn J. G. Effectiveness of breast cancer surveillance in BRCA1/2 gene mutation carriers and women with high familial risk. *J Clin Oncol*, 19(4):924–30, 2001.

[99] Travis L. B., Hill D. A., Dores G. M., Gospodarowicz M., van Leeuwen F. E., Holowaty E., Glimelius B., Andersson M., Wiklund T., Lynch C. F., Van't Veer M. B., Glimelius I., Storm H., Pukkala E., Stovall M., Curtis R., Boice, J. D. J., and Gilbert E. Breast cancer following radiotherapy and chemotherapy among young women with Hodgkin disease. *JAMA*, 290(4): 465–75, 2003.

[100] Degnim A. C., Visscher D. W., Berman H. K., Frost M. H., Sellers T. A., Vierkant R. A., Maloney S. D., Pankratz V. S., de Groen P. C., Lingle W. L., Ghosh K., Penheiter L., Tlsty T., Melton, L. J. r., Reynolds C. A., and Hartmann L. C. Stratification of breast cancer risk in women with atypia: a Mayo cohort study. *J Clin Oncol*, 25(19):2671–7, 2007.

[101] Bedrosian I., Mick R., Orel S. G., Schnall M., Reynolds C., Spitz F. R., Callans L. S., Buzby G. P., Rosato E. F., Fraker D. L., and Czerniecki B. J. Changes in the surgical management of patients with breast carcinoma based on preoperative magnetic resonance imaging. *Cancer*, 98(3):468–473, 2003.

[102] Warner E., Plewes D. B., Hill K. A., Causer P. A., Zubovits J. T., Jong R. A., Cutrara M. R., DeBoer G., Yaffe M. J., Messner S. J., Meschino W. S., Piron C. A., and Narod S. A. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *JAMA*, 292(11):1317–25, 2004.

[103] Rutgers E. J. and Tuut M. K. [CBO guideline 'Breast cancer: screening and diagnosis']. *Ned Tijdschr Geneeskd*, 145(3):115–9, 2001.

[104] Kolb T. M., Lichy J., and Newhouse J. H. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology*, 225(1):165–75, 2002.

[105] Dalmis M. U., Litjens G., Holland K., Setio A., Mann R., Karssemeijer N., and Gubern-Merida A. Using deep learning to segment breast and fibroglanduar tissue in MRI volumes. *Med Phys*, 44(2):533–546, 2016.

[106] ACR BI-RADS Atlas. (4), 2003.

[107] Newcombe R. G. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*, 17(8):857–72, 1998.

[108] Rebbeck T. R., Friebel T., Lynch H. T., Neuhausen S. L., van 't Veer L., Garber J. E., Evans G. R., Narod S. A., Isaacs C., Matloff E., Daly M. B., Olopade O. I., and Weber B. L. Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE study group. *J Clin Oncol*, 22(6):1055–62, 2004.

[109] Freitas V., Crystal P., Kulkarni S. R., Ghai S., Bukhanov K., Escallon J., and Scaranelo A. M. The value of breast MRI in high-risk patients with newly diagnosed breast cancer to exclude invasive disease in the contralateral prophylactic mastectomy: Is there a role to choose wisely patients for sentinel node biopsy? *Cancer Med*, 5(6):1031–1036, 2016.

[110] Boughey J. C., Khakpour N., Meric-Bernstam F., Ross M. I., Kuerer H. M., Singletary S. E., Babiera G. V., Arun B., Hunt K. K., and Bedrosian I. Selective use of sentinel lymph node surgery during prophylactic mastectomy. *Cancer*, 107(7):1440–7, 2006.

[111] Phi X. A., Saadatmand S., De Bock G. H., Warner E., Sardanelli F., Leach M. O., Riedl C. C., Trop I., Hooning M. J., Mandel R., Santoro F., Kwan-Lim G., Helbich T. H., Tilanus-Linthorst M. M., van den Heuvel E. R., and Houssami N. Contribution of mammography to MRI screening in BRCA mutation carriers by BRCA status and age: individual patient data meta-analysis. *Br J Cancer*, 114(6):631–7, 2016.

[112] Houssami N., Abraham L. A., Miglioretti D. L., Sickles E. A., Kerlikowske K., Buist D. S., Geller B. M., Muss H. B., and Irwig L. Accuracy and outcomes of screening mammography in women with a personal history of early-stage breast cancer. *JAMA*, 305(8):790–9, 2011.

[113] Obdeijn I. M., Winter-Warnars G. A., Mann R. M., Hooning M. J., Hunink M. G., and Tilanus-Linthorst M. M. Should we screen BRCA1 mutation carriers only with MRI? A multicenter study. *Breast Cancer Res Treat*, 144(3):577–82, 2014.

[114] Saadatmand S., Obdeijn I. M., Rutgers E. J., Oosterwijk J. C., Tollenaar R. A., Woldringh G. H., Bergers E., Verhoef C., Heijnsdijk E. A., Hooning M. J., de Koning H. J., and Tilanus-Linthorst M. M. Survival benefit in women with BRCA1 mutation or familial risk in the MRI screening study (MRISC). *Int J Cancer*, 137(7):1729–38, 2015.

[115] Evans D. G., Kesavan N., Lim Y., Gadde S., Hurley E., Massat N. J., Maxwell A. J., Ingham S., Eeles R., Leach M. O., Howell A., and Duffy S. W. MRI breast screening in high-risk women: cancer detection and survival analysis. *Breast Cancer Res Treat*, 145(3):663–72, 2014.

[116] Nystrom L., Bjurstam N., Jonsson H., Zackrisson S., and Frisell J. Reduced breast cancer mortality after 20+ years of follow-up in the Swedish randomized controlled mammography trials in Malmö, Stockholm, and Göteborg. *J Med Screen*, 2016.

[117] Ford D., Easton D. F., Stratton M., Narod S., Goldgar D., Devilee P., Bishop D. T., Weber B., Lenoir G., Chang-Claude J., Sobol H., Teare M. D., Struewing J., Arason A., Scherneck S., Peto J., Rebbeck T. R., Tonin P., Neuhausen S., Barkardottir R., Eyfjord J., Lynch H., Ponder B. A., Gayther S. A., Zelada-Hedman M., and et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet*, 62(3):676–89, 1998.

[118] Tilanus-Linthorst M. M., Obdeijn I. M., Bartels K. C., de Koning H. J., and Oudkerk M. First experiences in screening women at high risk for breast cancer with MR imaging. *Breast Cancer*

Res Treat, 63(1):53–60, 2000.

[119] Komenaka I. K., Ditkoff B. A., Joseph K. A., Russo D., Gorroochurn P., Ward M., Horowitz E., El-Tamer M. B., and Schnabel F. R. The development of interval breast malignancies in patients with BRCA mutations. *Cancer*, 100(10):2079–83, 2004.

[120] Plevritis S. K., Kurian A. W., Sigal B. M., Daniel B. L., Ikeda D. M., Stockdale F. E., and Garber A. M. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. *JAMA*, 295(20):2374–84, 2006.

[121] Schrading S. and Kuhl C. K. Mammographic, US, and MR imaging phenotypes of familial breast cancer. *Radiology*, 246(1):58–70, 2008.

[122] Tilanus-Linthorst M., Verhoog L., Obdeijn I. M., Bartels K., Menke-Pluymers M., Eggermont A., Klijn J., Meijers-Heijboer H., van der Kwast T., and Brekelmans C. A BRCA1/2 mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography. *Int J Cancer*, 102(1):91–5, 2002.

[123] Berrington de Gonzalez A., Berg C. D., Visvanathan K., and Robson M. Estimated risk of radiation-induced breast cancer from mammographic screening for young BRCA mutation carriers. *J Natl Cancer Inst*, 101(3):205–9, 2009.

[124] Andrieu N., Easton D. F., Chang-Claude J., Rookus M. A., Brohet R., Cardis E., Antoniou A. C., Wagner T., Simard J., Evans G., Peock S., Fricker J. P., Nogues C., Van't Veer L., Van Leeuwen F. E., and Goldgar D. E. Effect of chest X-rays on the risk of breast cancer among BRCA1/2 mutation carriers in the international BRCA1/2 carrier cohort study: a report from the EMBRACE, GENEPSO, GEO-HEBON, and IBCCS Collaborators' Group. *J Clin Oncol*, 24(21):3361–6, 2006.

[125] Vreemann S., Gubern-Merida A., Schlooz-Vries M. S., Bult P., van Gils C. H., Hoogerbrugge N., Karssemeijer N., and Mann R. M. Influence of risk category and screening round on the performance of an MR Imaging and mammography screening program in carriers of the BRCA mutation and other women at increased risk. *Radiology*, 286(2), 2018.

[126] Molleran V. and Mahoney M. C. The BI-RADS breast magnetic resonance imaging lexicon. *Magn Reson Imaging Clin N Am*, 18(2):171–85, vii, 2010.

[127] Metcalfe K. A., Lubinski J., Ghadirian P., Lynch H., Kim-Sing C., Friedman E., Foulkes W. D., Domchek S., Ainsworth P., Isaacs C., Tung N., Gronwald J., Cummings S., Wagner T., Manoukian S., Moller P., Weitzel J., Sun P., and Narod S. A. Predictors of contralateral prophylactic mastectomy in women with a BRCA1 or BRCA2 mutation: the Hereditary Breast Cancer Clinical Study Group. *J Clin Oncol*, 26(7):1093–7, 2008.

[128] Sung J. S., Stamler S., Brooks J., Kaplan J., Huang T., Dershaw D. D., Lee C. H., Morris E. A., and Comstock C. E. Breast cancers detected at screening MR Imaging and mammography in patients at high risk: Method of detection reflects tumor histopathologic results. *Radiology*, 280 (3):716–22, 2016.

[129] Lord S. J., Lei W., Craft P., Cawson J. N., Morris I., Walleser S., Griffiths A., Parker S., and Houssami N. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *Eur J Cancer*, 43(13):1905–17, 2007.

[130] Heijnsdijk E. A., Warner E., Gilbert F. J., Tilanus-Linthorst M. M., Evans G., Causer P. A., Eeles

R. A., Kaas R., Draisma G., Ramsay E. A., Warren R. M., Hill K. A., Hoogerbrugge N., Wasser M. N., Bergers E., Oosterwijk J. C., Hooning M. J., Rutgers E. J., Klijn J. G., Plewes D. B., Leach M. O., and de Koning H. J. Differences in natural history between breast cancers in BRCA1 and BRCA2 mutation carriers and effects of MRI screening-MRISC, MARIBS, and Canadian studies combined. *Cancer Epidemiol Biomarkers Prev*, 21(9):1458–68, 2012.

[131] Narayan A. K., Visvanathan K., and Harvey S. C. Comparative effectiveness of breast MRI and mammography in screening young women with elevated risk of developing breast cancer: a retrospective cohort study. *Breast Cancer Res Treat*, 158(3):583–9, 2016.

[132] Gubern-Merida A., Vreemann S., Marti R., Melendez J., Lardenoije S., Mann R. M., Karssemeijer N., and Platel B. Automated detection of breast cancer in false-negative screening MRI studies from women at increased risk. *Eur J Radiol*, 85(2):472–9, 2016.

[133] Berg W. A., Zhang Z., Lehrer D., Jong R. A., Pisano E. D., Barr R. G., Bohm-Velez M., Mahoney M. C., Evans, W. P. r., Larsen L. H., Morton M. J., Mendelson E. B., Farria D. M., Cormack J. B., Marques H. S., Adams A., Yeh N. M., and Gabrielli G. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA*, 307(13):1394–404, 2012.

[134] Lang K., Andersson I., and Zackrisson S. Breast cancer detection in digital breast tomosynthesis and digital mammography - a side-by-side review of discrepant cases. *Br J Radiol*, 87(1040): 20140080, 2014.

[135] Kuhl C. K. Abbreviated breast MRI for screening women with dense breast: The EA1141 Trial. *Br J Radiol*, page 20170441, 2017.

[136] Bleyer A. and Welch H. G. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med*, 367(21):1998–2005, 2012.

[137] Puliti D., Duffy S. W., Miccinesi G., de Koning H., Lynge E., Zappa M., and Paci E. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen*, 19 Suppl 1:42–56, 2012.

[138] Krammer J., Pinker-Domenig K., Robson M. E., Gonen M., Bernard-Davila B., Morris E. A., Mangino D. A., and Jochelson M. S. Breast cancer detection and tumor characteristics in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res Treat*, 163(3):565–571, 2017.

[139] Martin J. E., Moskowitz M., and Milbrath J. R. Breast cancer missed by mammography. *AJR Am J Roentgenol*, 132(5):737–9, 1979.

[140] Broeders M. J., Onland-Moret N. C., Rijken H. J., Hendriks J. H., Verbeek A. L., and Holland R. Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection. *Eur J Cancer*, 39(12):1770–5, 2003.

[141] Weber R. J., van Bommel R. M., Louwman M. W., Nederend J., Voogd A. C., Jansen F. H., Tjan-Heijnen V. C., and Duijm L. E. Characteristics and prognosis of interval cancers after biennial screen-film or full-field digital screening mammography. *Breast Cancer Res Treat*, 158(3):471–83, 2016.

[142] Destounis S. V., DiNitto P., Logan-Young W., Bonaccio E., Zuley M. L., and Willison K. M. Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? Initial experience. *Radiology*, 232(2):578–84, 2004.

[143] Thurfjell E. L., Lernevall K. A., and Taube A. A. Benefit of independent double reading in a population-based mammography screening program. *Radiology*, 191(1):241–4, 1994.

[144] Anttinen I., Pamilo M., Soiva M., and Roiha M. Double reading of mammography screening films–one radiologist or two? *Clin Radiol*, 48(6):414–21, 1993.

[145] Yamaguchi K., Schacht D., Newstead G. M., Bradbury A. R., Verp M. S., Olopade O. I., and Abe H. Breast cancer detected on an incident (second or subsequent) round of screening MRI: MRI features of false-negative cases. *AJR Am J Roentgenol*, 201(5):1155–63, 2013.

[146] Pages E. B., Millet I., Hoa D., Doyon F. C., and Taourel P. Undiagnosed breast cancer at MR imaging: analysis of causes. *Radiology*, 264(1):40–50, 2012.

[147] Ritter F., Boskamp T., Homeyer A., Laue H., Schwier M., Link F., and Peitgen H. O. Medical image analysis. *IEEE Pulse*, 2(6):60–70, 2011.

[148] Gubern-Merida A., Marti R., Melendez J., Hauth J. L., Mann R. M., Karssemeijer N., and Platel B. Automated localization of breast cancer in DCE-MRI. *Med Image Anal*, 20(1):265–74, 2015.

[149] Liberman L., Morris E. A., Lee M. J., Kaplan J. B., LaTrenta L. R., Menell J. H., Abramson A. F., Dashnaw S. M., Ballon D. J., and Dershaw D. D. Breast lesions detected on MR imaging: features and positive predictive value. *AJR Am J Roentgenol*, 179(1):171–8, 2002.

[150] Vinokur A. D., Threatt B. A., Caplan R. D., and Zimmerman B. L. Physical and psychosocial functioning and adjustment to breast cancer. Long-term follow-up of a screening population. *Cancer*, 63(2):394–405, 1989.

[151] Kirsh V. A., Chiarelli A. M., Edwards S. A., O'Malley F. P., Shumak R. S., Yaffe M. J., and Boyd N. F. Tumor characteristics associated with mammographic detection of breast cancer in the Ontario breast screening program. *J Natl Cancer Inst*, 103(12):942–50, 2011.

[152] Shah P. K., Austin J. H., White C. S., Patel P., Haramati L. B., Pearson G. D., Shiau M. C., and Berkmen Y. M. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology*, 226(1):235–41, 2003.

[153] Turkington P. M., Kennan N., and Greenstone M. A. Misinterpretation of the chest X-ray as a factor in the delayed diagnosis of lung cancer. *Postgrad Med J*, 78(917):158–60, 2002.

[154] Tilanus-Linthorst M. M., Obdeijn I. M., Hop W. C., Causer P. A., Leach M. O., Warner E., Pointon L., Hill K., Klijn J. G., Warren R. M., and Gilbert F. J. BRCA1 mutation and young age predict fast breast cancer growth in the Dutch, United Kingdom, and Canadian magnetic resonance imaging screening trials. *Clin Cancer Res*, 13(24):7357–62, 2007.

[155] Shimauchi A., Jansen S. A., Abe H., Jaskowiak N., Schmidt R. A., and Newstead G. M. Breast cancers not detected at MRI: review of false-negative lesions. *AJR Am J Roentgenol*, 194(6):1674–9, 2010.

[156] Gordon P. B., Borugian M. J., and Warren Burhenne L. J. A true screening environment for review of interval breast cancers: pilot study to reduce bias. *Radiology*, 245(2):411–5, 2007.

[157] Baek J. E., Kim S. H., and Lee A. W. Background parenchymal enhancement in breast MRIs of breast cancer patients: impact on tumor size estimation. *Eur J Radiol*, 83(8):1356–62, 2014.

[158] Kuhl C. The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology*, 244(2):356–78, 2007.

[159] DeMartini W. B., Liu F., Peacock S., Eby P. R., Gutierrez R. L., and Lehman C. D. Background parenchymal enhancement on breast MRI: impact on diagnostic performance. *AJR Am J Roentgenol*, 198(4):W373–80, 2012.

[160] Uematsu T., Kasami M., and Watanabe J. Background enhancement of mammary glandular tissue on breast dynamic MRI: imaging features and effect on assessment of breast cancer extent. *Breast Cancer*, 19(3):259–65, 2012.

[161] Morris E. A. Diagnostic breast MR imaging: current status and future directions. *Magn Reson Imaging Clin N Am*, 18(1):57–74, 2010.

[162] Telegrafo M., Rella L., Stabile Ianora A. A., Angelelli G., and Moschetta M. Effect of background parenchymal enhancement on breast cancer detection with magnetic resonance imaging. *Diagn Interv Imaging*, 97(3):315–20, 2016.

[163] Park S. Y., Kang D. K., and Kim T. H. Does background parenchymal enhancement on MRI affect the rate of positive resection margin in breast cancer patients? *Br J Radiol*, 88(1046): 20140638, 2015.

[164] Giess C. S., Yeh E. D., Raza S., and Birdwell R. L. Background parenchymal enhancement at breast MR imaging: normal patterns, diagnostic challenges, and potential for false-positive and false-negative interpretation. *Radiographics*, 34(1):234–47, 2014.

[165] Saadatmand S., Tilanus-Linthorst M. M., Rutgers E. J., Hoogerbrugge N., Oosterwijk J. C., Tollenaar R. A., Hooning M., Loo C. E., Obdeijn I. M., Heijnsdijk E. A., and de Koning H. J. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst*, 105(17):1314–21, 2013.

[166] Feig S. Cost-effectiveness of mammography, MRI, and ultrasonography for breast cancer screening. *Radiol Clin North Am*, 48(5):879–91, 2010.

[167] Elston C. W. and Ellis I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, 19(5):403–10, 1991.

[168] van der Velden B. H., Dmitriev I., Loo C. E., Pijnappel R. M., and Gilhuijs K. G. Association between parenchymal enhancement of the contralateral breast in dynamic contrast-enhanced MR Imaging and outcome of patients with unilateral invasive breast cancer. *Radiology*, 276(3): 675–85, 2015.

[169] Kim M. Y., Choi N., Yang J. H., Yoo Y. B., and Park K. S. Background parenchymal enhancement on breast MRI and mammographic breast density: correlation with tumour characteristics. *Clin Radiol*, 70(7):706–10, 2015.

[170] Mohsin S. K., Weiss H., Havighurst T., Clark G. M., Berardo M., Roanh le D., To T. V., Qian Z., Love R. R., and Allred D. C. Progesterone receptor by immunohistochemistry and clinical outcome in breast cancer: a validation study. *Mod Pathol*, 17(12):1545–54, 2004.

[171] Harvey J. M., Clark G. M., Osborne C. K., and Allred D. C. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol*, 17(5):1474–81, 1999.

[172] van Zelst J. C. M., Balkenhol M., Tan T., Rutten M., Imhof-Tas M., Bult P., Karssemeijer N., and Mann R. M. Sonographic phenotypes of molecular subtypes of invasive ductal cancer in

automated 3-D breast ultrasound. *Ultrasound Med Biol*, 43(9):1820–1828, 2017.

[173] Telegrafo M., Rella L., Stabile Ianora A. A., Angelelli G., and Moschetta M. Breast MRI background parenchymal enhancement (BPE) correlates with the risk of breast cancer. *Magn Reson Imaging*, 34(2):173–6, 2016.

[174] Hambly N. M., Liberman L., Dershaw D. D., Brennan S., and Morris E. A. Background parenchymal enhancement on baseline screening breast MRI: impact on biopsy rate and short-interval follow-up. *AJR Am J Roentgenol*, 196(1):218–24, 2011.

[175] Bulut N. and Altundag K. Does estrogen receptor determination affect prognosis in early stage breast cancers? *Int J Clin Exp Med*, 8(11):21454–9, 2015.

[176] Delille J. P., Slanetz P. J., Yeh E. D., Kopans D. B., Halpern E. F., and Garrido L. Hormone replacement therapy in postmenopausal women: breast tissue perfusion determined with MR imaging – initial observations. *Radiology*, 235(1):36–41, 2005.

[177] Ellis R. L. Optimal timing of breast MRI examinations for premenopausal women who do not have a normal menstrual cycle. *AJR Am J Roentgenol*, 193(6):1738–40, 2009.

[178] Ross R. K., Paganini-Hill A., Wan P. C., and Pike M. C. Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. *J Natl Cancer Inst*, 92(4):328–32, 2000.

[179] Schairer C., Lubin J., Troisi R., Sturgeon S., Brinton L., and Hoover R. Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer risk. *JAMA*, 283(4):485–91, 2000.

[180] Colditz G. A. Relationship between estrogen levels, use of hormone replacement therapy, and breast cancer. *J Natl Cancer Inst*, 90(11):814–23, 1998.

[181] Chlebowski R. T., Hendrix S. L., Langer R. D., Stefanick M. L., Gass M., Lane D., Rodabough R. J., Gilligan M. A., Cyr M. G., Thomson C. A., Khandekar J., Petrovitch H., and McTiernan A. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. *JAMA*, 289(24):3243–53, 2003.

[182] Magnusson C., Baron J. A., Correia N., Bergstrom R., Adami H. O., and Persson I. Breast-cancer risk following long-term oestrogen- and oestrogen-progestin-replacement therapy. *Int J Cancer*, 81(3):339–44, 1999.

[183] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet*, 350(9084):1047–59, 1997.

[184] Clemons M. and Goss P. Estrogen and the risk of breast cancer. *N Engl J Med*, 344(4):276–85, 2001.

[185] Tabar L., Fagerberg G., Chen H. H., Duffy S. W., and Gad A. Tumour development, histology and grade of breast cancers: prognosis and progression. *Int J Cancer*, 66(4):413–9, 1996.

[186] Moelans C. B., de Wegers R. A., Monsuurs H. N., Maess A. H., and van Diest P. J. Molecular differences between ductal carcinoma in situ and adjacent invasive breast carcinoma: a multiplex ligation-dependent probe amplification study. *Cell Oncol (Dordr)*, 34(5):475–82, 2011.

[187] Hu Y. BRCA1, hormone, and tissue-specific tumor suppression. *Int J Biol Sci*, 5(1):20–7, 2009.

[188] Vinnicombe S. How I report breast magnetic resonance imaging studies for breast cancer staging and screening. *Cancer Imaging*, 16(1):17, 2016.

[189] Millet I., Pages E., Hoa D., Merigeaud S., Curros Doyon F., Prat X., and Taourel P. Pearls and pitfalls in breast MRI. *Br J Radiol*, 85(1011):197–207, 2012.

[190] Erbas B., Amos A., Fletcher A., Kavanagh A. M., and Gertig D. M. Incidence of invasive breast cancer and ductal carcinoma in situ in a screening program by age: should older women continue screening? *Cancer Epidemiol Biomarkers Prev*, 13(10):1569–73, 2004.

[191] Uematsu T., Kasami M., Yuen S., Igarashi T., and Nasu H. Comparison of 3- and 1.5-T dynamic breast MRI for visualization of spiculated masses previously identified using mammography. *AJR Am J Roentgenol*, 198(6):W611–7, 2012.

[192] Domchek S. M., Eisen A., Calzone K., Stopfer J., Blackwood A., and Weber B. L. Application of breast cancer risk prediction models in clinical practice. *J Clin Oncol*, 21(4):593–601, 2003.

[193] Cintolo-Gonzalez J., Braun D., Blackford A., Mazzola E., Acar A., Plichta J., Griffin M., and Hughes K. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast Cancer Res Treat*, 164(2):263–284, 2017.

[194] Holm J., Humphreys K., Li J., Ploner A., Cheddad A., Eriksson M., Tornberg S., Hall P., and Czene K. Risk factors and tumor characteristics of interval cancers by mammographic density. *J Clin Oncol*, 33(9):1030–7, 2015.

[195] Mitchell G., Antoniou A. C., Warren R., Peock S., Brown J., Davies R., Mattison J., Cook M., Warsi I., Evans D. G., Eccles D., Douglas F., Paterson J., Hodgson S., Izatt L., Cole T., Burgess L., Eeles R., and Easton D. F. Mammographic density and breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Cancer Res*, 66(3):1866–72, 2006.

[196] Bertrand K. A., Scott C. G., Tamimi R. M., Jensen M. R., Pankratz V. S., Norman A. D., Visscher D. W., Couch F. J., Shepherd J., Chen Y. Y., Fan B., Wu F. F., Ma L., Beck A. H., Cummings S. R., Kerlikowske K., and Vachon C. M. Dense and nondense mammographic area and risk of breast cancer by age and tumor characteristics. *Cancer Epidemiol Biomarkers Prev*, 24(5):798–809, 2015.

[197] Bertrand K. A., Tamimi R. M., Scott C. G., Jensen M. R., Pankratz V., Visscher D., Norman A., Couch F., Shepherd J., Fan B., Chen Y. Y., Ma L., Beck A. H., Cummings S. R., Kerlikowske K., and Vachon C. M. Mammographic density and risk of breast cancer by age and tumor characteristics. *Breast Cancer Res*, 15(6):R104, 2013.

[198] Ding J., Warren R., Girling A., Thompson D., and Easton D. Mammographic density, estrogen receptor status and other breast cancer tumor characteristics. *Breast J*, 16(3):279–89, 2010.

[199] Eng A., Gallant Z., Shepherd J., McCormack V., Li J., Dowsett M., Vinnicombe S., Allen S., and dos Santos-Silva I. Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. *Breast Cancer Res*, 16(5):439, 2014.

[200] Pike M. C. and Pearce C. L. Mammographic density, MRI background parenchymal enhancement and breast cancer risk. *Ann Oncol*, 24 Suppl 8:viii37–viii41, 2013.

[201] Morris E. A. Diagnostic breast MR imaging: current status and future directions. *Radiol Clin North Am*, 45(5):863–80, vii, 2007.

[202] Dalmis M., Gubern-Merida A., Borelli C., Vreemann S., Mann R., and Karssemeijer N. A fully automated system for quantification of background parenchymal enhancement in breast DCE-MRI. In *SPIE*, volume 9785, 2016.

[203] Steyerberg E. Clinical Prediction Models. page 500. Springer-Verlag New York, 2009.

[204] Wanders J. O., Holland K., Veldhuis W. B., Mann R. M., Pijnappel R. M., Peeters P. H., van Gils C. H., and Karssemeijer N. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat*, 162(1):95–103, 2017.

[205] Passaperuma K., Warner E., Hill K. A., Gunasekara A., and Yaffe M. J. Is mammographic breast density a breast cancer risk factor in women with BRCA mutations? *J Clin Oncol*, 28(23):3779–83, 2010.

[206] Gubern-Merida A., Kallenberg M., Mann R. M., Marti R., and Karssemeijer N. Breast segmentation and density estimation in breast MRI: a fully automatic framework. *IEEE J Biomed Health Inform*, 19(1):349–57, 2015.

[207] Bennani-Baiti B., Dietzel M., and Baltzer P. A. MRI background parenchymal enhancement is not associated with breast cancer. *PLoS One*, 11(7):e0158573, 2016.

[208] Kuhl C. K., Bieling H. B., Gieseke J., Kreft B. P., Sommer T., Lutterbey G., and Schild H. H. Healthy premenopausal breast parenchyma in dynamic contrast-enhanced MR imaging of the breast: normal contrast medium enhancement and cyclical-phase dependency. *Radiology*, 203 (1):137–44, 1997.

[209] King V., Kaplan J., Pike M. C., Liberman L., David Dershaw D., Lee C. H., Brooks J. D., and Morris E. A. Impact of tamoxifen on amount of fibroglandular tissue, background parenchymal enhancement, and cysts on breast magnetic resonance imaging. *Breast J*, 18(6):527–34, 2012.

[210] Widschwendter M., Rosenthal A. N., Philpott S., Rizzuto I., Fraser L., Hayward J., Intermaggio M. P., Edlund C. K., Ramus S. J., Gayther S. A., Dubeau L., Fourkala E. O., Zaikin A., Menon U., and Jacobs I. J. The sex hormone system in carriers of BRCA1/2 mutations: a case-control study. *Lancet Oncol*, 14(12):1226–32, 2013.

[211] Wu S., Weinstein S. P., DeLeo, M. J. r., Conant E. F., Chen J., Domchek S. M., and Kontos D. Quantitative assessment of background parenchymal enhancement in breast MRI predicts response to risk-reducing salpingo-oophorectomy: preliminary evaluation in a cohort of BRCA1/2 mutation carriers. *Breast Cancer Res*, 17:67, 2015.

[212] Giess C. S., Raza S., and Birdwell R. L. Patterns of nonmasslike enhancement at screening breast MR imaging of high-risk premenopausal women. *Radiographics*, 33(5):1343–60, 2013.

[213] Brennan S. B., Sung J. S., Dershaw D. D., Liberman L., and Morris E. A. Cancellation of MR imaging-guided breast biopsy due to lesion nonvisualization: frequency and follow-up. *Radiology*, 261(1):92–9, 2011.

[214] van der Velden B. H. M., Elias S. G., Bismeijer T., Loo C. E., Viergever M. A., Wessels L. F. A., and Gilhuijs K. G. A. Complementary value of contralateral parenchymal enhancement on DCE-MRI to prognostic models and molecular assays in high-risk ER(+)/HER2(-) breast cancer. *Clin Cancer Res*, 23(21):6505–6515, 2017.

[215] Saadatmand S., Bretveld R., Siesling S., and Tilanus-Linthorst M. M. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173,797

patients. *BMJ*, 351:h4901, 2015.

[216] Paap E., Verbeek A. L., Botterweck A. A., van Doorne-Nagtegaal H. J., Imhof-Tas M., de Koning H. J., Otto S. J., de Munck L., van der Steen A., Holland R., den Heeten G. J., and Broeders M. J. Breast cancer screening halves the risk of breast cancer death: a case-referent study. *Breast*, 23 (4):439–44, 2014.

[217] Kaiser W. A. and Zeitler E. MR imaging of the breast: fast imaging sequences with and without Gd-DTPA. preliminary observations. *Radiology*, 170(3 Pt 1):681–6, 1989.

[218] Brekelmans C. T., Tilanus-Linthorst M. M., Seynaeve C., vd Ouweland A., Menke-Pluymers M. B., Bartels C. C., Kriege M., van Geel A. N., Burger C. W., Eggermont A. M., Meijers-Heijboer H., and Klijn J. G. Tumour characteristics, survival and prognostic factors of hereditary breast cancer from BRCA2-, BRCA1- and non-BRCA1/2 families as compared to sporadic breast cancer cases. *Eur J Cancer*, 43(5):867–76, 2007.

[219] Taneja C., Edelsberg J., Weycker D., Guo A., Oster G., and Weinreb J. Cost effectiveness of breast cancer screening with contrast-enhanced MRI in high-risk women. *J Am Coll Radiol*, 6(3): 171–9, 2009.

[220] Griebsch I., Brown J., Boggis C., Dixon A., Dixon M., Easton D., Eeles R., Evans D. G., Gilbert F. J., Hawnaur J., Kessar P., Lakhani S. R., Moss S. M., Nerurkar A., Padhani A. R., Pointon L. J., Potterton J., Thompson D., Turnbull L. W., Walker L. G., Warren R., Leach M. O., and Group U. K. M. R. I. i. B. S. S. Cost-effectiveness of screening with contrast enhanced magnetic resonance imaging vs X-ray mammography of women at a high familial risk of breast cancer. *Br J Cancer*, 95(7):801–10, 2006.

[221] Kuhl C. K., Strobel K., Bieling H., Leutner C., Schild H. H., and Schrading S. Supplemental breast MR Imaging screening of women with average risk of breast cancer. *Radiology*, 283(2): 361–370, 2017.

[222] Saadatmand S., Rutgers E. J., Tollenaar R. A., Zonderland H. M., Ausems M. G., Keymeulen K. B., Schlooz-Vries M. S., Koppert L. B., Heijnsdijk E. A., Seynaeve C., Verhoef C., Oosterwijk J. C., Obdeijn I. M., de Koning H. J., and Tilanus-Linthorst M. M. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicentre randomized controlled trial. *BMC Cancer*, 12:440, 2012.

[223] Chen S. Q., Huang M., Shen Y. Y., Liu C. L., and Xu C. X. Application of abbreviated protocol of magnetic resonance imaging for breast cancer screening in dense breast tissue. *Acad Radiol*, 24(3):316–320, 2017.

[224] Chen S. Q., Huang M., Shen Y. Y., Liu C. L., and Xu C. X. Abbreviated MRI protocols for detecting breast cancer in women with dense breasts. *Korean J Radiol*, 18(3):470–475, 2017.

[225] Romeo V., Cuocolo R., Liuzzi R., Riccardi A., Accurso A., Acquaviva A., Buonocore R., and Imbriaco M. Preliminary results of a simplified breast mri protocol to characterize breast lesions: Comparison with a full diagnostic protocol and a review of the current literature. *Acad Radiol*, 24(11):1387–1394, 2017.

[226] Strahle D. A., Pathak D. R., Sierra A., Saha S., Strahle C., and Devisetty K. Systematic development of an abbreviated protocol for screening breast magnetic resonance imaging. *Breast Cancer Res Treat*, 162(2):283–295, 2017.

[227] Tofts P. S., Berkowitz B., and Schnall M. D. Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. *Magn Reson Med*, 33(4):564–8, 1995.

[228] Veltman J., Stoutjesdijk M., Mann R., Huisman H. J., Barentsz J. O., Blickman J. G., and Boetes C. Contrast-enhanced magnetic resonance imaging of the breast: the value of pharmacokinetic parameters derived from fast dynamic imaging during initial enhancement in classifying lesions. *Eur Radiol*, 18(6):1123–33, 2008.

[229] Le Y., Kipfer H., Majidi S., Holz S., Dale B., Geppert C., Kroeker R., and Lin C. Application of time-resolved angiography with stochastic trajectories (TWIST)-Dixon in dynamic contrast-enhanced (DCE) breast MRI. *J Magn Reson Imaging*, 38(5):1033–42, 2013.

[230] Saranathan M., Rettmann D. W., Hargreaves B. A., Lipson J. A., and Daniel B. L. Variable spatiotemporal resolution three-dimensional Dixon sequence for rapid dynamic contrast-enhanced breast MRI. *J Magn Reson Imaging*, 40(6):1392–9, 2014.

[231] Mus R. D., Borelli C., Bult P., Weiland E., Karssemeijer N., Barentsz J. O., Gubern-Merida A., Platel B., and Mann R. M. Time to enhancement derived from ultrafast breast MRI as a novel parameter to discriminate benign from malignant breast lesions. *Eur J Radiol*, 89:90–96, 2017.

[232] Onishi N., Kataoka M., Kanao S., Sagawa H., Iima M., Nickel M. D., Toi M., and Togashi K. Ultrafast dynamic contrast-enhanced MRI of the breast using compressed sensing: breast cancer diagnosis based on separate visualization of breast arteries and veins. *J Magn Reson Imaging*, 47 (1):97–104, 2018.

[233] Platel B., Mus R., Welte T., Karssemeijer N., and Mann R. Automated characterization of breast lesions imaged with an ultrafast DCE-MR protocol. *IEEE Trans Med Imaging*, 33(2):225–32, 2014.

[234] Dalmis M. U., Gubern-Merida A., Vreemann S., Karssemeijer N., Mann R., and Platel B. A computer-aided diagnosis system for breast DCE-MRI at high spatiotemporal resolution. *Med Phys*, 43(1):84, 2016.

[235] Benkert T., Block K. T., Heller S., Moccaldi M., Sodickson D. K., Kim S. G., and Moy L. Comprehensive dynamic contrast-enhanced 3D magnetic resonance imaging of the breast with fat/water separation and high spatiotemporal resolution using radial sampling, compressed sensing, and parallel imaging. *Invest Radiol*, 52(10):583–589, 2017.

[236] Vreemann S., Rodriguez-Ruiz A., Nickel D., Heacock L., Appelman L., van Zelst J., Karssemeijer N., Weiland E., Maas M., Moy L., Kiefer B., and Mann R. M. Compressed sensing for breast MRI: Resolving the trade-off between spatial and temporal resolution. *Invest Radiol*, 2017.

[237] Petrillo A., Fusco R., Sansone M., Cerbone M., Filice S., Porto A., Rubulotta M. R., D'Aiuto M., Avino F., Di Bonito M., and Botti G. Abbreviated breast dynamic contrast-enhanced MR imaging for lesion detection and characterization: the experience of an Italian oncologic center. *Breast Cancer Res Treat*, 164(2):401–410, 2017.

[238] Panigrahi B., Mullen L., Falomo E., Panigrahi B., and Harvey S. An abbreviated protocol for high-risk screening breast magnetic resonance imaging: Impact on performance metrics and bi-rads assessment. *Acad Radiol*, 24(9):1132–1138, 2017.

[239] Peters N. H., Borel Rinkes I. H., Zuithoff N. P., Mali W. P., Moons K. G., and Peeters P. H. Meta-analysis of MR imaging in the diagnosis of breast lesions. *Radiology*, 246(1):116–24, 2008.

[240] Piccoli C. W. Contrast-enhanced breast MRI: factors affecting sensitivity and specificity. *Eur*

Radiol, 7 Suppl 5:281–8, 1997.

[241] Partridge S. C., Rahbar H., Murthy R., Chai X., Kurland B. F., DeMartini W. B., and Lehman C. D. Improved diagnostic accuracy of breast MRI through combined apparent diffusion coefficients and dynamic contrast-enhanced kinetics. *Magn Reson Med*, 65(6):1759–67, 2011.

[242] Hargreaves B. A., Saranathan M., Sung K., and Daniel B. L. Accelerated breast MRI with compressed sensing. *Eur J Radiol*, 81 Suppl 1:S54–5, 2012.

[243] Nickel D., Chen X., Mailhe B., Wang Q., Son Y., Min Lee J., and Kiefer B. Motion-resolved 3D dynamic contrast enhanced liver MRI. *Proc. Intl. Soc. Mag. Reson. Med.*, 24, 2016.

[244] van Ginneken B., van Rikxoort E., Lafebre S., Jacobs C., Schmidt M., Kuhnigk J.-M., Prokop M., Schaefer-Prokop C., Charbonnier J.-P., Hogeweg L., Maduskar P., Gallardo Estrella L., Philipsen R., and Lassen B. CIRRUS Lung: An optimized workflow for quantitative image analysis of thoracic computed tomography and chest radiography or major pulmonary diseases - chronic obstructive pulmonary disease, lung cancer and tuberculosis. In *Proc. RSNA*, Chicago, IL, 2013.

[245] Boetes C., Barentsz J. O., Mus R. D., van der Sluis R. F., van Erning L. J., Hendriks J. H., Holland R., and Ruys S. H. MR characterization of suspicious breast lesions with a gadolinium-enhanced TurboFLASH subtraction technique. *Radiology*, 193(3):777–81, 1994.

[246] Owen H. W., Dockerty M. B., and Gray H. K. Occult carcinoma of the breast. *Surg Gynecol Obstet*, 98(3):302–8, 1954.

[247] Kuijs V. J., Moossdorff M., Schipper R. J., Beets-Tan R. G., Heuts E. M., Keymeulen K. B., Smidt M. L., and Lobbes M. B. The role of MRI in axillary lymph node imaging in breast cancer patients: a systematic review. *Insights Imaging*, 6(2):203–15, 2015.

[248] Lee S., Yi A., Chang J. M., Cho N., and Moon W. K. Should the axilla be included in screening ultrasound? *Proc. RSNA*, 2016.

[249] Huang W., Tudorica L. A., Li X., Thakur S. B., Chen Y., Morris E. A., Tagge I. J., Korenblit M. E., Rooney W. D., Koutcher J. A., and Springer, C. S. J. Discrimination of benign and malignant breast lesions by using shutter-speed dynamic contrast-enhanced MR imaging. *Radiology*, 261 (2):394–403, 2011.

[250] Rapacchi S., Natsuaki Y., Plotnik A., Gabriel S., Laub G., Finn J. P., and Hu P. Reducing view-sharing using compressed sensing in time-resolved contrast-enhanced magnetic resonance angiography. *Magn Reson Med*, 74(2):474–81, 2015.

[251] Rueckert D., Sonoda L. I., Hayes C., Hill D. L., Leach M. O., and Hawkes D. J. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*, 18(8):712–21, 1999.

[252] Stoutjesdijk M. J., Futterer J. J., Boetes C., van Die L. E., Jager G., and Barentsz J. O. Variability in the description of morphologic and contrast enhancement characteristics of breast lesions on magnetic resonance imaging. *Invest Radiol*, 40(6):355–62, 2005.

[253] Miller B. A., Feuer E. J., and Hankey B. F. Recent incidence trends for breast cancer in women and the relevance of early detection: an update. *CA Cancer J Clin*, 43(1):27–41, 1993.

[254] Chu K. C., Tarone R. E., Kessler L. G., Ries L. A., Hankey B. F., Miller B. A., and Edwards B. K. Recent trends in U.S. breast cancer incidence, survival, and mortality rates. *J Natl Cancer Inst*,

88(21):1571–9, 1996.

[255] Plevritis S., Munoz D., Kurian A., Stout N., Alagoz O., Near A., Lee S., van den Broek J., Huang X., Schechter C., Sprague B., Song J., de Koning H., Trentham-Dietz A., van Ravesteyn N., Gangnon R., Chandler Y., Li Y., Xu C., Ergun M., Huang H., Berry D., and Mandelblatt J. Association of screening and treatment with breast cancer mortality by molecular subtype in US women, 2000-2012. *JAMA*, 319(2):154–164, 2018.

[256] Zackrisson S., Andersson I., Janzon L., Manjer J., and Garne J. P. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ*, 332 (7543):689–92, 2006.

[257] Svendsen A. L., Olsen A. H., von Euler-Chelpin M., and Lynge E. Breast cancer incidence after the introduction of mammography screening: What should be expected? *Cancer*, 106(9): 1883–90, 2006.

[258] Jorgensen K. J. and Gotzsche P. C. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ*, 339:b2587, 2009.

[259] Duffy S. W., Agbaje O., Tabar L., Vitak B., Bjurstam N., Bjorneld L., Myles J. P., and Warwick J. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res*, 7(6):258–65, 2005.

[260] Tamaki K., Ishida T., Miyashita M., Amari M., Ohuchi N., Tamaki N., and Sasano H. Correlation between mammographic findings and corresponding histopathology: potential predictors for biological characteristics of breast diseases. *Cancer Sci*, 102(12):2179–85, 2011.

[261] Hoff S. R., Abrahamsen A. L., Samset J. H., Vigeland E., Klepp O., and Hofvind S. Breast cancer: missed interval and screening-detected cancer at full-field digital mammography and screen-film mammography – results from a retrospective review. *Radiology*, 264(2):378–86, 2012.

[262] Manning D., Ethell S., and Donovan T. Categories of observer error from eye tracking and AFROC data. *Proc. SPIE*, 90, 2004.

[263] Pataky R., Armstrong L., Chia S., Coldman A. J., Kim-Sing C., McGillivray B., Scott J., Wilson C. M., and Peacock S. Cost-effectiveness of MRI for breast cancer screening in BRCA1/2 mutation carriers. *BMC Cancer*, 13:339, 2013.

[264] Ehteshami Bejnordi B., Veta M., van Diest P., van Ginneken B., Karssemijer N., Litjens G., van der Laak A., and "the CAMELYON16 Consortium". Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*, 318(22):2199–2210, 2017.

Dankwoord

Graag wil ik een aantal mensen bedanken voor het mede mogelijk maken van dit proefschrift. Zonder al deze mensen was dit proefschrift er niet geweest en daarvoor wil ik jullie allen bedanken, ook de mensen die ik vergeet bij naam te noemen en dat zullen er vast een aantal zijn. Alvast mijn excuses hiervoor, het zijn de zenuwen.

Allereerst, Ritse, zonder jou was het niet gelukt. De discussies over welke statistische testen gebruikt zouden moeten worden waren altijd erg nuttig, ook al was ik af en toe eigenwijs genoeg om het dan toch nog steeds op mijn eigen manier te doen. Je enthousiasme en ideeën hielden me altijd op de been als ik het even niet zag zitten. Heel erg bedankt!

Nico, bedankt dat ik binnen de mamma-groep mocht promoveren en dat je me vrij liet in wat ik samen met Ritse wilde uitzoeken. Vaak was je benieuwd waar we nu weer mee bezig waren en dan planden we een afspraak in om het daarover, maar ook over de verdere planning, te hebben. Ik heb altijd erg genoten van deze gesprekken en deze hielpen me de structuur te vinden binnen mijn promotietraject.

Albert, although I see you more as a friend, thank you for being my co-promotor. Our collaboration started with the so-called prior paper. Originally, the idea was that I was going to submit that one, and you would wait until that one was published to submit your CAD version. However, I was so slow that you already published yours before I even had a first draft version of mine. I always enjoyed working with you, thank you!

Bram P, heel erg bedankt dat je er was om me te helpen tijdens het begin van mijn promotietraject. Ik vond het in het begin lastig om de goede keuzes te maken, maar jij was er altijd om me op weg te helpen. Helaas ging het slechter met je gezondheid en moest je je daarop focussen. Ik heb erg genoten van de samenwerking en hoop je snel nog eens te zien!

Ik wil graag alle co-auteurs bedanken, zonder jullie vragen, opmerkingen en hulp waren de artikelen nooit van dit niveau geweest.

Ik wil alle leden van de beoordelingscommissie bedanken. Bedankt voor het beoordelen van dit proefschrift en ik kijk uit naar de discussies tijdens de verdediging.

Mehmet, thank you for working together with me on one of the topics that I really loved to do: BPE. Without you, it was unthinkable that we could perform this study

on such a large dataset.

Colin, Jan-Jurre, Rick, en Albert, ik wil jullie bedanken voor de heerlijke discussies tijdens de koffiemomentjes aan het begin van de ochtend. Ik vond het fijn dat er altijd mensen waren die naar mijn gezeur wilden luisteren.

Ik wil iedereen van DIAG bedanken het was altijd gezellig. In het bijzonder wil ik de mensen bij wie ik op de kamer heb gezeten bedanken, maar omdat ik zo vaak verplaatst ben is dat toch zeker half DIAG, dus bij dezen (in willekeurige volgorde): Freerk, Kaman, Katharina, Sjoerd, Paul, James, Thijs, Wendelien, Anton, Gabriel, Saurabh, Joris, Arnaud, Sil, Aya (Ajay) en Midas, bedankt! Freerk, we begonnen samen op -1 en zo heb jij mijn ups en downs van dichtbij meegemaakt. Bedankt dat je altijd voor me klaar stond. And of course, Alejandro. I want to thank you for everything, you were always there to help me when I needed it the most (although you were always too busy). Your help and you listening to my complaints helped me through the end of my PhD. Babak, thank you for your contribution to science[264]. Jan, we waren een goed team. Je hielp me altijd aan de nieuwste roddels en daarnaast zal ik de bezoekjes aan de McDonalds tijdens de ECR nooit vergeten.

Daarnaast wil ik ook de onderzoekers van de MRS-groep en klinische onderzoeksgroep bedanken, maar in het bijzonder Linda en Marnix. Ik voelde me gelijk in de groep opgenomen toen we samen op congres gingen. Linda, bedankt voor de steun en het luisterende oor. Marnix, bedankt voor de hulp met alle MRI en statistiek vragen. Solange, Charlotte en Marjolein, bedankt voor het helpen met van alles. And of course, Gerrita and the BT III 4.0 team from Siemens, thank you for your support and the educational experience.

Marloes, Marjolein, Lisanne en Charlotte, de Chica's, het was altijd gezellig met jullie en ik hoop dat we met onze drukke levens toch nog tijd kunnen maken voor onze filmavondjes en vakanties. Misschien wordt het binnenkort maar weer eens tijd om af te spreken.

Stefan, bedankt dat je er was voor de koffiemomentjes waarin je me wijze raad gaf. Iedereen heeft die af en toe nodig, en ik al helemaal.

Lisanne en Ellen, oud-huisgenootjes, bedankt voor de fijne tijd in Nijmegen, ik vond het altijd gezellig. En mijn andere oud-huisgenootje uit Enschede, Marleen, bedankt voor het maken van de voorkant en natuurlijk het cat-sitten.

Jolien, bedankt dat je een van mijn beste vriendinnen bent en dat ik tijdens het laatste half jaar van mijn PhD af en toe mijn frustratie kon uiten tijdens het schilderen van de babykamer. Samen met DJ zijn we het beste team ooit!

Merel, Sjoerdje, Freek, Eelke, Donald, Isabelle en Bram, jullie zijn fantastische vrienden op wie ik kan rekenen, ook als het even tegenzit. Maar vooral als we op een feestje staan is het altijd gezellig.

Kaman en Jan-Jurre, dankjulliewel dat jullie mijn paranimfen willen zijn. Kaman, wij begonnen samen op -1 en samen hebben we prachtige avonturen beleefd, van karaokeavonden tot ritjes met de rolstoel, het is een eindeloze lijst, het was altijd gezellig. Je bent een fantastische vriendin! En Jurre, jij stond ook altijd voor me klaar, met je grappige verhalen, het snoep, en natuurlijk de bezoekjes aan de Fest op vrijdag. De kapsalons zullen nooit meer hetzelfde zijn zonder jou.

Fleur en Karlijn, sestra's, bedankt dat jullie me af en toe van het werk hielden. Karlijn met weetjes over je nieuwe huis en Fleur met de opdrachten in SPSS. Karlijn, bedankt voor de goede zorgen voor Henk, ik hoop dat je nog een keer wilt oppassen.

Papa en mama, dank voor de steun die jullie me altijd gegeven hebben. Papa, ik hoop dat je na het lezen van de Nederlandse samenvatting nu begrijpt wat ik al die jaren heb uitgevoerd.

Lieve Pieter, zonder jou had ik het niet gered. Ik kan me niet voorstellen hoe jij je af en toe moet hebben gevoeld als ik weer terneergeslagen thuiskwam omdat mijn paper voor de zoveelste keer was afgewezen en ik het gevoel had dat ik terug bij af was. Je probeert me nog steeds te leren kitesurfen, maar na die gebroken voet weet ik niet zeker of ik het ooit ga leren. Bedankt dat je altijd voor me klaar staat! Je bent een schat.

Curriculum Vitae

Suzan Vreemann was born in Neede, the Netherlands, on July 29th, 1989. In 2007, she started the Technical Medicine program at the University of Twente, in Enschede, the Netherlands. Her master thesis entitled "A novel magnetic based approach for lymph node detection" was a collaboration between the University of Twente and the Medisch Spectrum Twente, Enschede. In February 2014, Suzan started as a PhD student in the Diagnostic Image Analysis Group at the Department of Radiology and Nuclear Medicine, Radboudumc in Nijmegen, the Netherlands. She started focusing on the integrated classification of breast cancers in the EU-funded project VPH-PRISM but also gained knowledge about the breast cancer screening for women at high risk of breast cancer. The results she carried out at the Diagnostic Image Analysis Group are described in this thesis. In June 2018, she started working at DICA (Dutch Institute for Clinical Auditing) as a team leader.